Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

DISSERTATION THESIS

Long-range cross-correlations: Tests, estimators and applications

Author: **PhDr. Ladislav Krištoufek**

Supervisor: **Mgr. Lukáš Vácha, Ph.D.**

Academic Year: **2012/2013**

Declaration of Authorship

Hereby I declare that this thesis has been compiled independently, using the listed resources and literature only.

Prague, July 23, 2013

Signature

Acknowledgments

I would like to express gratitude mainly to my supervisor Lukáš Vácha and Jozef Baruník for their motivation and frequent discussions which have lead to this thesis. Without their constructive and critical attitude, the thesis would have never crystalized into this form. I would also like to thank prof. Miloslav Vošvrda for his support and understanding in following rather unorthodox paths in economics and finance. My thanks also go to fellow Ph.D. students attending the research seminars Nonlinear Dynamic Economic Systems whose comments have helped in my research progress as well.

I am also thankful for comments and discussions at conferences that helped during my research path significantly. The research leading to the final version of the dissertation was presented and discussed at International Conference on Computational and Financial Econometrics 2011 & 2012, International Conference on Computing in Economics and Finance 2010, 2011 & 2012, Econophysics Colloquium 2010, 2011 & 2012, WEHIA 2010 & 2012, International Conference on Econophysics 2011, International Conference on Statistical Physics 2011 and International Conference on Mathematical Methods in Economics 2010, 2011 & 2012.

Following the research path would never be possible without support, love, patience and understanding of my wife and my family as a whole. To Lord God, gratitude goes for his love, forgiveness and mercy without which none of this would be possible.

The thesis has been written with funding from GAUK grants nos. 118310 and 1110213, and GAČR grants nos. 402/09/0965 and 402/09/H045. The support is well appreciated.

Abstract

The motivation of this thesis is to provide a basic framework for treating long-range cross-correlated processes while keeping the methodology and assumptions as general as possible. Starting from the definition of long-range cross-correlated processes as jointly stationary processes with asymptotically power-law decaying cross-correlation function, we show that such definition implies a divergent at origin cross-power spectrum and power-law scaling of covariances of partial sums of the long-range cross-correlated processes. Chapter 2 describes these and other basic definitions and propositions together with necessary proofs. Chapter 3 then introduces several processes which possess long-range cross-correlated series properties. Apart from cases when the memory parameter of the bivariate memory is a simple average of the parameters of the separate processes, we also introduce a new kind of process, which we call the mixed-correlated ARFIMA, which allows to control for both the bivariate and univariate memory parameters. Chapter 4 deals with tests for a presence of long-range cross-correlations. We develop three new tests, and Monte-Carlo-simulation-based statistical power and size of the tests are compared. The newly introduced tests strongly surpass the already existing one. In Chapter 5, we cover the estimators of long-range cross-correlation parameter of choice – the bivariate Hurst exponent. The estimators are split into two groups based on their domain of operation – time and frequency. In addition to four already existing estimators, one of which has been introduced by the author of this thesis, we introduce two new estimators. As another novelty, we reconfigure the estimators so that the power law coherency can be estimated as well. Finite sample statistical properties (bias, variance and mean squared error) of the estimators are compared for various specifications. In Chapter 6, we analyze the leverage effect between financial returns and volatility from a perspective of the long-range cross-correlations. We then conclude and hint several challenges for further research.

JEL Classification	C14, C15, C51, G17
Keywords	cross-correlations, long-term memory, power law coherency, Hurst exponent

Author's e-mail	kristoufek@ies-prague.org
Supervisor's e-mail	vachal@utia.cas.cz

Abstrakt

Hlavní motivací dizertační práce je navrhnout základní rámec pro analýzu dlouhé paměti v křížových korelacích v co nejobecnější podobě. Přes definici procesů s dlouhou pamětí v křížových korelacích jako sdruženě stacionárních s asymptoticky mocninně klesající funkcí křížových korelací ukazujeme, že tato definice implikuje v počátku divergentní křížové spektrum a mocninný zákon kovariancí parciálních sum křížově persistentních procesů. V kapitole 2 popisujeme tyto a další definice a věty společně s potřebnými důkazy. Kapitola 3 zavádí několik procesů, které lze popsat jako procesy s dlouhočasovanými křížovými korelacemi. Kromě případů, kdy je parametr dvourozměrné dlouhé paměti průměrem parametrů jednotlivých procesů, také zavádíme nový typ procesu, který nazýváme jako smíšenně-korelované ARFIMA procesy, u kterých lze manipulovat parametry paměti nejen u jednotlivých procesů, ale i u křížové paměti. Kapitola 4 diskutuje testy na přítomnost dlouhé paměti v křížových korelacích. Zavádíme tři nové testy a srovnáváme je s již existujícím testem a na základě Monte Carlo simulací ukazujeme, že nové testy silně dominují již existující test. V kapitole 5 pokrýváme odhady parametrů dlouhé paměti mezi dvěma procesy – bivariátního Hurstova exponentu. Odhady jsou rozděleny do dvou skupin podle definic na časové a frekvenční odhady. Kromě čtyř již existujících odhadů, kde jeden byl představen autorem této práce, zavádíme dva další. Nově také předefinováváme odhady, aby byly schopné odhadovat mocninou koherenci. Statistické vlastnosti odhadů pro konečné řady (vychýlení, rozptyl a střední čtvercová chyba) jsou poravnány pro různé specifikace modelů. V kapitole 6 analyzujeme pákový efekt mezi finančními výnosy a volatilitou z pohledu dlouhočasových křížových korelací. V závěru se zaměřujeme na možné směrování dalšího výzkumu.

Klasifikace JEL	C14, C15, C51, G17
Klíčová slova	křížové korelace, dlouhá paměť', mocninný zákon koherence, Hurstův exponent

E-mail autora	`kristoufek@ies-prague.org`
E-mail vedoucího práce	`vachal@utia.cas.cz`

Contents

List of Tables

List of Figures

Chapter 1

Introduction

A notion of long-range dependence has become a stable part of financial econometrics in recent years as several financial series have been shown to possess statistical properties connected with long-range correlations. Even though the long-range correlations have been initially studied in the water flows of the Nile River in the pioneering work of Hurst (1951), the research has quickly spread to other branches of science including finance and financial econometrics. The early studies are mainly due to Benoît Mandelbrot and his colleagues (Mandelbrot 1966; 1967; Mandelbrot & Wallis 1968; Mandelbrot & van Ness 1968; Mandelbrot 1971) who later focused more on the multifractality of the series, which is tightly connected to the long-range dependence (Mandelbrot *et al.* 1997; Mandelbrot 1999; 2005; Calvet & Fisher 2008). In the financial econometrics literature, the long-range correlations are standardly studied as fractionally differenced series, which were pioneered by Granger & Joyeux (1980) and Hosking (1981). Fractional differencing implies that the process has an infinite memory, i.e. that a realization of a process at a specific time is influenced by all its previous realizations. Strength of the memory is described by parameter d, which is standardly bounded between $-0.5 \leq d < 0.5$ for stationary processes. A persistent behavior of the series characterized by $d > 0$ has been found in several financial series – realized volatility, range-based volatility, absolute and squared returns, and traded volume (Bollerslev & Jubinski 1999; Thomakos & Wang 2003; Poon & Granger 2003; 2005; Chen *et al.* 2006; Forsberg & Ghysels 2007; Fleming & Kirby 2011). In words, the persistence of the series implies local trending while the stationarity still holds, which evidently brings a potential for modeling and mainly forecasting. This has lead to several new models capturing such property, most importantly to a fractional

generalization of standard ARMA processes – ARFIMA processes (Granger & Joyeux 1980) – and fractional generalization of standard GARCH processes – FIGARCH processes (Baillie *et al.* 1996; Bollerslev & Mikkelsen 1996) – and derived models. The research on the fractional differencing and integration has also given rise to a fractional cointegration technique, which is a generalization of the original cointegration applied frequently in economics and finance (Cheung & Lai 1993; Baillie & Bollerslev 1994). This way, the long-term equilibrium relationship is not limited only to a unit root series but it covers wide range of fractionally integrated processes (Gil-Alana & Hualde 2009).

In recent years, there has been a growing interest in generalization of the long-range dependence concept for higher dimensions. In most cases, the multivariate setting is utilized for obtaining more efficient estimators of the fractional differencing parameters d or Hurst exponents H of the separate processes. This approach leads mainly to spectrum-based estimators which assume a specific case of the cross-power spectrum – namely divergent at origin cross-power spectrum possibly with some other specifics. Robinson (1995b) uses GPH (originally by Geweke & Porter-Hudak (1983) for univariate series) for estimation of d_1 and d_2 of the separate processes. Lobato (1997) estimates the separate fractional differencing parameters using the average periodogram estimator (APE) assuming the processes are persistent ($d_1, d_2 > 0$). Lobato (1999) further estimates the same parameters using the Gaussian semi-parametric estimator (GSE). Nielsen (2004a) introduces a maximum likelihood estimator for several scenarios in the multivariate setting. Shimotsu (2007) estimates the set of parameters d_1 and d_2 using the GSE but additionally takes the phase of the cross-power spectrum into consideration. Nielsen (2011) uses the phase information as well and proposes the local Whittle estimator of separate d parameters of the multivariate series which also covers the non-stationary values of parameter $d \in \langle 0.5, +\infty)$. The methods are nicely summarized and compared in Sela (2010). All the above mentioned estimators are used to estimate the memory parameters of the separate processes. Sela & Hurvich (2012) are, to our best knowledge, the first ones in this branch of estimators to propose a fractional differencing parameter estimator between two processes – d_{12} – generalizing the APE using the cross-power spectrum assumptions. Sela & Hurvich (2012) also discuss the possibility of the bivariate parameter d_{12} not being the average of d_1 and d_2, the possibility noted by Lobato (1997), in more detail introducing a concept of a power law coherency and anti-cointegration.

The frequency domain approach to the estimation usually assumes a rela-

tively specific form of the underlying multivariate process. There are several types of processes, which have been proposed in the literature, possessing the desirable properties of the divergent at origin cross-power spectrum. Apart from obvious generalizations of ARFIMA processes into the multivariate setting (Shimotsu 2007; Tsay 2010; Nielsen 2011), less standard processes have been introduced as well (e.g. fractionally integrated continuous time autoregressive moving average processes of Marquardt (2007)). Fractional Brownian motion has been also generalized into the multivariate setting (Lavancier *et al.* 2009; Amblard & Coeurjolly 2011; Coeurjolly *et al.* 2012; Amblard *et al.* 2012). All these models have a built-in equivalency between d_{12} and the average of the separate fractional differencing parameters d_1 and d_2. As it is desirable to be able to manipulate with d_{12} irrespective (or at least partially irrespective) of d_1 and d_2, there are types of processes proposed in the literature which enable such a manipulation. Sela & Hurvich (2012) propose a very general setting of infinite-order moving average models which allow to control for parameters d_{12}, d_1 and d_2 as long as $d_{12} \leq \frac{1}{2}(d_1 + d_2)$. Podobnik *et al.* (2008) introduce a two-component ARFIMA processes, which are based on ARFIMA-like mixing of two processes. Unfortunately, the authors neither provide any clue how to control the bivariate parameter H_{xy} nor is it evident whether the processes are even stationary.

Assuming a specific model obviously has both advantages and disadvantages. On one hand, there is the possibility to derive asymptotic properties such as bias, consistency, variance and distribution of the estimators in most cases. On the other hand, the estimators can be highly biased and inconsistent if the underlying process differs from the assumed one. Avoiding the assumptions of the underlying processes, there are time domain estimators which assume only a specific form of the cross-correlation function – namely an asymptotically hyperbolically, or power-law, decaying one. Such approach has the exactly opposite advantages and disadvantages to the model-assuming spectrum-based estimators, asymptotic statistical properties are hard or impossible to show but the approach is more general and can cover a wider range of models. These estimators are usually labeled as the heuristic ones because the asymptotic properties, as already noted, cannot be derived and the estimates are usually taken only as approximate due to a relatively high variance of the estimators.

The time domain estimators are usually based on a power law scaling of various covariance measures of partial sums of the processes. Podobnik & Stanley

(2008) propose the detrended cross-correlation analysis (DCCA) as a bivariate generalization of the detrended fluctuation analysis (DFA) of Peng *et al.* (1994). Kristoufek (2011) generalizes the height-height correlation analysis of Barabasi *et al.* (1991) and Barabasi & Vicsek (1991), and the generalized Hurst exponent approach of Di Matteo *et al.* (2003) into the multifractal height cross-correlation analysis (MF-HXA). And the detrending moving average (DMA) of Alessio *et al.* (2002) is generalized by He & Chen (2011a) forming the detrended moving-average cross-correlation analysis (DMCA). These methodologies are used to estimate the bivariate Hurst exponent H_{xy}, which is connected to the bivariate fractional differencing parameter through $H_{xy} = d_{12} + 0.5$. The time domain estimators thus focus mainly on the bivariate memory parameter and they provide no efficiency gain over the univariate time domain estimators for the separate memory parameters.

Out of these four time and frequency domain estimators of H_{xy} and d_{12}, only Sela & Hurvich (2012) provide a finite sample study for their averaged periodogram estimator. Statistical performance of the other three – time domain – estimators has not been provided yet. These estimators are the bivariate generalizations of the methods with finite sample properties which have been quite frequently discussed in the literature, and it has been shown that their variance can be very high for short samples (for extensive Monte Carlo studies, see Taqqu *et al.* (1995); Taqqu & Teverovsky (1996); Barunik & Kristoufek (2010); Kristoufek (2010c)). Moreover, most of the univariate estimators are rather sensitive to a presence of the short-term memory in the underlying process and even a moderate short-range dependence can be easily mistaken for the long-range dependence. Such issues have been also quite frequently discussed in the literature (Lo 1991; Teverovsky *et al.* 1999; Barunik *et al.* 2012). It is very likely that the bivariate estimators suffer from the same issues. However, none of these have been discussed for the long-range cross-correlated processes yet. More severely, as the power-law decay of the cross-correlation function has been usually treated as an automatic property of the multivariate long-memory models, the concept has not been given a proper focus and consideration yet (apart from Sela (2010); Sela & Hurvich (2012)).

The motivation of this thesis is to provide the essential framework for treating long-range cross-correlated processes while remaining as general as possible. Starting from the definition of the long-range cross-correlated processes as jointly stationary processes with asymptotically power-law decaying cross-correlation function, we follow the steps of the standard univariate long-range

dependence texts such as Beran (1994) and Samorodnitsky (2006) to show that such definition implies a divergent at origin magnitude of the cross-power spectrum and a power-law scaling of covariances of partial sums of the long-range cross-correlated processes. Chapter 2 describes these and other basic definitions and propositions together with necessary proofs. Chapter 3 introduces several processes which possess the long-range cross-correlated series properties. Apart from the cases when the memory parameter of the bivariate memory is a simple average of the parameters of the separate processes, we also introduce a new kind of process which allows to control both the bivariate and univariate memory parameters and we call it the mixed-correlated ARFIMA. Chapter 4 deals with tests for a presence of the long-range cross-correlations. We develop three new tests and Monte-Carlo-simulation-based statistical powers and sizes of the tests are compared. The newly introduced tests strongly surpass the already existing one. In Chapter 5, we cover the estimators of the long-range cross-correlation parameter of choice – the bivariate Hurst exponent. The estimators are split into two groups based on their domain of operation – time and frequency. Apart from four already existing estimators, one of which has been introduced by the author of this thesis in Kristoufek (2011), we present two new frequency-based estimators. As another novelty, we reconfigure the estimators so that the power law coherency can be estimated as well. Extensive Monte Carlo study is supplemented and basic finite sample statistical properties (bias, variance and mean squared error) of the estimators are compared for various specifications. In Chapter 6, we apply the proposed methodology to analyze the leverage effect between financial returns and volatility. The dissertation thesis tries to provide a very first coherent text on long-range cross-correlations and proposes a way of how to statistically treat them. Obviously, the text is not a complete statement on the long-range cross-correlations and is naturally limited due to a very general definition of the long-term memory between two series. The main motivation remains in initiating the discussion on the topic since most of the literature utilizes the multivariate setting only for efficiency gains in the univariate parameters estimation. Several open questions remain and these are partly discussed in the conclusion of this thesis.

The text is structured in a way that only the most important aspects are included in the main text. A derivation of the cross-correlation structures and cross-power spectra of the processes of Chapter 3 are included in Appendix A. Monte-Carlo-based sizes and powers of the tests introduced in Chapter 4 are summarized in tables in Appendix B. Finite sample properties of the estimators

of the bivariate Hurst exponent and the power law coherency are covered in Appendix C. Additional figures and tables for the leverage effect analysis are supplied in Appendix D. To allow for possible replication of the results, all necessary R-project codes are available at `http://ies.fsv.cuni.cz/en/staff/kristoufek`.

Chapter 2

Long-range cross-correlated processes

In this chapter, we introduce the basic statistical concepts used in the following text. The theory concerning the univariate long-range dependence is only shortly reviewed while the main focus is put on definitions and propositions introducing a concept of the long-range cross-correlations. The long-range cross-correlations are defined through a hyperbolic decay of cross-correlation function and we show that this definition implies, among others, a divergent at origin magnitude of the cross-power spectrum and a power-law scaling of partial sums of the processes. These are later used as a basis for estimators of the bivariate Hurst exponent.

In the text, we are mainly interested in the asymptotic properties and as the properties are often in a form of a power law, we are also usually interested just in the proportional relationship. For these purposes, we use the following notation – "$\approx$" stands for "approximately", "$\propto$" is used as "proportional to", and "$\sim$" means "asymptotically proportional to" or "asymptotically approximately proportional to" depending on a context. To avoid confusion, we use terms "auto-correlation", "cross-correlation" and "correlation" in the following way. The first one is used for serial correlation of a single series. The second one is used for the dependence between one series and lagged/led values of the other series. And the third term is used to measure dependence between two series with no leads or lags.

2.1 Preliminaries

We work with stochastic processes as a family of random variables $X(\omega, t)$ where $\omega \in \Omega$ is an outcome out of a sample space Ω, and t belongs to an index set and in general, $t = 1, \ldots, T$ where T is a number of realizations of the process. For a better legibility, we use $X_t \equiv X(\omega, t)$ and $\{x_t\}$ as a series or realizations of a random variable X_t throughout the text.

Cumulative distribution function $F_{X_t}(x) = P(\omega : X_t \leq x)$ gives very basic information about the statistical properties of the random variable X_t and mainly about its moments. Using the notation $\langle \bullet \rangle$ as an expectation operator, we are mainly interested in moments described by functions

$$\mu_t = \langle X_t \rangle \tag{2.1}$$

$$\sigma_t^2 = \langle (X_t - \mu_t)^2 \rangle = \langle X_t^2 \rangle - \langle \mu_t \rangle^2 \tag{2.2}$$

$$\gamma(t_1, t_2) = \langle (X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2}) \rangle = \langle (X_{t_1} X_{t_2}) \rangle - \langle X_{t_1} \rangle \langle X_{t_2} \rangle \tag{2.3}$$

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sqrt{\sigma_{t_1}^2 \sigma_{t_2}^2}} \tag{2.4}$$

where μ_t is a mean function, σ_t^2 is a variance function, $\gamma(t_1, t_2)$ is a covariance function between random variables X_{t_1} and X_{t_2}, and $\rho(t_1, t_2)$ is a correlation function between random variables X_{t_1} and X_{t_2}.

Stationarity of processes is one of the main assumptions of a majority of definitions and propositions we present in the text. For this purpose, we consider wide-sense stationary processes in the following meaning:

Definition 2.1 (Wide-sense stationarity). Stochastic process $\{x_t\}$ is said to be wide-sense stationary if $\mu_t = \mu$ and $\sigma_t^2 = \sigma^2$ are constant, and $\gamma(t_1, t_2) = \gamma(k)$ and $\rho(t_1, t_2) = \rho(k)$ are functions of the time lag $k \in \mathbb{Z}$ and are independent of t. Moreover, we have $\langle |X_t| \rangle < +\infty$ and $\langle X_t^2 \rangle < +\infty$.

Wide-sense stationary (WSS) processes are also called weakly or covariance stationary, see e.g. Wei (2006) for discussion. In words, the wide-sense stationary processes have constant mean and variance and their correlation structure depends on a lag k and not on time t or a starting point t_0. We utilize wide-sense stationary processes as the second-order wide-sense stationary processes so that the higher moments are not considered, which might be limiting in some cases but makes it applicable to an applied work. If not stated otherwise, we refer to WSS processes in Definition 2.1 as simply "stationary".

In Definition 2.1, the covariance function $\gamma(k)$ and the correlation function $\rho(k)$ are standardly labeled as the auto-covariance and auto-correlation functions, respectively. Specifically for stationary processes with a constant mean μ and a constant variance σ^2, we have $\gamma(k) = \langle (X_t - \mu)(X_{t+k} - \mu) \rangle$ and $\rho(k) = \frac{\gamma(k)}{\sigma^2} = \frac{\gamma(k)}{\gamma(0)}$. Both functions $\gamma(k)$ and $\rho(k)$ are even and thus symmetric around $k = 0$ and positively semidefinite (Wei 2006).

The auto-correlation properties of the stochastic processes can be also analyzed from a spectral perspective. Let $\{x_t\}$ be a real-valued stationary process with absolutely summable auto-covariance function so that $\sum_{k=-\infty}^{+\infty} |\gamma(k)| < +\infty$. Power spectrum $f(\lambda)$ with a frequency $-\pi \leq \lambda \leq \pi$ is defined as

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma(k) \exp(-ik\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma(k) \cos(k\lambda) \propto \sum_{k=-\infty}^{+\infty} \rho(k) \cos(k\lambda). \tag{2.5}$$

Through the inverse Fourier transform, we can obtain the auto-covariance and auto-correlation functions back as

$$\gamma(k) = \int_{-\pi}^{\pi} f(\lambda) \exp(ik\lambda) d\lambda \tag{2.6}$$

$$\rho(k) = \frac{1}{\sigma^2} \int_{-\pi}^{\pi} f(\lambda) \exp(ik\lambda) d\lambda \tag{2.7}$$

so that there is no information loss. The power spectrum $f(\lambda)$ is a continuous real-valued non-negative function so that $f(|\lambda|) = f(\lambda) = |f(\lambda)|$, i.e. it suffices to analyze the power spectrum only for $0 < \lambda \leq \pi$. It is simply a Fourier transform of the auto-covariance function and can thus be seen as a decomposition of the variance of the process to specific frequencies λ.

Let's now consider two stochastic processes $\{x_t\}$ and $\{y_t\}$ formed of random variables X_t and Y_t with the identical time index set. Apart from the above mentioned mean, variance, covariance and correlation functions, we further define

$$\gamma_{XY}(t_1, t_2) = \langle (X_{t_1} - \mu_{X,t_1})(Y_{t_2} - \mu_{Y,t_2}) \rangle = \langle (X_{t_1} Y_{t_2}) \rangle - \langle X_{t_1} \rangle \langle Y_{t_2} \rangle \tag{2.8}$$

$$\rho_{XY}(t_1, t_2) = \frac{\gamma_{XY}(t_1, t_2)}{\sqrt{\sigma^2_{X,t_1} \sigma^2_{Y,t_2}}} \tag{2.9}$$

as the cross-covariance and cross-correlation functions, respectively, between random variables X_t and Y_t.

Similarly to the wide-sense stationary processes, we also define the jointly wide-sense stationary processes:

Definition 2.2 (Joint wide-sense stationarity). Two stochastic processes $\{x_t\}$ and $\{y_t\}$ are said to be jointly wide-sense stationary if they are both wide-sense stationary and if it further holds that $\gamma_{XY}(t_1, t_2) = \gamma_{XY}(k)$ and $\rho_{XY}(t_1, t_2) = \rho_{XY}(k)$ are functions of the time lag $k \in \mathbb{Z}$ and are independent of t.

Therefore, for two processes to be jointly wide-sense stationary (J-WSS), they need to be separately wide-sense stationary and to have the cross-covariance and cross-correlation functions being dependent on a lag k and independent of any specific time position t or a starting point t_0. For J-WSS processes with absolutely summable cross-covariance function, i.e. $\sum_{k=-\infty}^{+\infty} |\gamma_{XY}(k)| < +\infty$, we define the cross-power spectrum for $-\pi \leq \lambda \leq \pi$ as

$$f_{XY}(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_{XY}(k) \exp(-ik\lambda) \propto \sum_{k=-\infty}^{+\infty} \rho_{XY}(k) \exp(-ik\lambda). \tag{2.10}$$

As the cross-covariance and cross-correlation functions are not necessarily (and most of the time they are not) symmetric, the cross-power spectrum is generally a complex-valued function. As the exponentials in Equation 2.10 form complex conjugates for $k < 0$ and $k > 0$, it again suffices to analyze only the case of $0 < \lambda \leq \pi$. The cross-covariance and cross-correlation functions can be obtained back from the cross-power spectrum in the same way as for the univariate case using the inverse Fourier transform as

$$\gamma_{XY}(k) = \int_{-\pi}^{\pi} f_{XY}(\lambda) \exp(ik\lambda) d\lambda \tag{2.11}$$

$$\rho_{XY}(k) = \frac{1}{\sigma_X \sigma_Y} \int_{-\pi}^{\pi} f_{XY}(\lambda) \exp(ik\lambda) d\lambda. \tag{2.12}$$

This provides the essential framework for treating the long-range cross-correlated processes. The next section starts with a general definition of such processes and derives several implications and scaling laws.

2.2 Scaling laws for long-range cross-correlated processes

The notion of the long-range dependence has a quite long history traditionally starting with Harold E. Hurst and his examination of the Nile River water flows (Hurst 1951). However, it has been Benoît Mandelbrot who is usually labelled as the founding father of the long-range dependence and a notion of fractality in statistics (Mandelbrot 1966; 1967; Mandelbrot & Wallis 1968; Mandelbrot & van Ness 1968; Mandelbrot 1971; 1972).

The definitions of the long-range dependence, or long-term memory, usually slightly vary across literature, see e.g. Beran (1994) and Samorodnitsky (2006) for reviews. The long-term memory is standardly associated with three phenomena – a hyperbolic (power-law) decay of the auto-correlation function of the process, a power-law scaling of some variance measure of the integrated series, and a diverging at the origin power spectrum. The connecting point to the scaling laws is the long-memory parameter – Hurst exponent H – which takes values between $0 \leq H < 1$ for stationary and invertible processes. For serially uncorrelated and short-range dependent processes, it holds that $H = 0.5$. Values above 0.5 indicate that the process is persistent and is reminiscent of a locally trending series which remains stationary. For values of H below 0.5, the process is anti-persistent and switches the sign of the realizations more frequently than a random process would. All these phenomena are well described in practically any text dealing with the long-range dependence (Beran 1994; Samorodnitsky 2006; Kantelhardt 2009; Giraitis *et al.* 2009) and we refer the reader to the references noted earlier for the detailed descriptions. Note that the notions we introduce later in the text for the bivariate series mostly hold for the univariate series as well.

2.2.1 Cross-correlation function scaling

In the same way as for the long-range dependence, there are several ways of how to define long-range cross-correlations – via scaling of the cross-correlation function or a slowly at infinity varying function, through a non-summability of the cross-correlation function, and a divergent at origin cross-power spectrum to name the most important ones. For our purposes, we stick with a definition connecting the long-range cross-correlations to the asymptotic power-law decay of the cross-correlation function to keep the definition as general as possible. To

be able to distinguish between the short-range and long-range cross-correlated processes, we present the definition of both, which are applied and used later in the text. In this case, we keep to the separation of the two types of processes as for the univariate case as presented in Kantelhardt (2009).

Definition 2.3 (Short-range cross-correlated processes). Two jointly stationary processes $\{x_t\}$ and $\{y_t\}$ are short-range cross-correlated if for $n > 0$ and/or $n < 0$, the cross-correlation function behaves as

$$\rho_{xy}(n) \propto \exp(-n/\delta) \tag{2.13}$$

with a characteristic time decay $0 \leq \delta < +\infty$.

Definition 2.4 (Long-range cross-correlated processes). Two jointly stationary processes $\{x_t\}$ and $\{y_t\}$ are long-range cross-correlated if for $n \to +\infty$, the cross-correlation function behaves as

$$\rho_{xy}(n) \propto n^{-\gamma_{xy}} \tag{2.14}$$

with a long-term memory parameter $0 < \gamma_{xy} < 1$.

It is thus sufficient to have at least one half of the cross-correlation function following one of the definitions to obtain short-range or long-range cross-correlated processes. Separation between the two types is thus strict – the short-range cross-correlated processes (which also encompass the pairwise uncorrelated but also pairwise correlated with no cross-correlations structure) possess an exponentially vanishing cross-correlation function while the long-range cross-correlated processes are characterized by an asymptotical power-law (or also hyperbolic) decaying cross-correlation function. Exponential decay is characteristic by several non-zero cross-correlations at low lags which quickly vanish to zero whereas the hyperbolic decay is connected to non-zero values of the cross-correlation function even for very high lags. Evidently, it is more complicated to analyze the asymptotic properties of the cross-correlation function for the latter type of processes so that frequently, the properties are studied on the scaling behavior of the cross-power spectrum or the covariance of the partial sums. These two approaches are discussed in the next sections of this chapter.

In contrast to the univariate case, the cross-correlation function is in general asymmetric. This creates several possibilities and combinations of be-

havior of the function such as one part being characteristic of the long-range cross-correlated processes and the other of the uncorrelated (or short-range cross-correlated) processes. However, for purposes of the long-range cross-correlations analysis, it is sufficient to assume that the series $\{x_t\}$ and $\{y_t\}$ are ordered in the manner of Definition 2.4. As a starting point, we usually assume that the decay of cross-correlations for the long-range cross-correlated processes is symmetric, i.e. $\rho_{xy}(n) \propto \rho_{xy}(-n) \propto n^{-\gamma_{xy}}$ as $n \to +\infty$, if not stated otherwise. In the case the decay is asymmetric with different scaling exponents γ^1_{xy} and γ^2_{xy}, we again assume that the series $\{x_t\}$ and $\{y_t\}$ are ordered so that $\rho_{xy}(n) \propto n^{-\min[\gamma^1_{xy},\gamma^2_{xy}]} \equiv n^{-\gamma_{xy}}$. It is evident later (specifically in Chapter 3) that the asymptotic behavior is dominated by the slower decay so that only $\min[\gamma^1_{xy}, \gamma^2_{xy}]$ is needed.

To keep the terminology parallel with the univariate long-range dependence, we call the processes according to Equation 2.14 as either long-range (long-term) cross-correlated or cross-persistent. If the cross-correlations are positive (negative), we call the processes positively (negatively) long-range (long-term) cross-correlated or positively (negatively) cross-persistent. Separation of the cross-persistent processes into either positively or negatively correlated is new compared to the univariate setting and must not be confused with the anti-persistent behavior. A bivariate generalization of the anti-persistent processes – cross-anti-persistence – is left out of the current text as the properties and implications of these processes are rather different from the cross-persistent ones and their inclusion would make the text hard to bear.

One of the alternative definitions of the long-range dependent processes is via non-summability of serial correlations (Beran 1994; Samorodnitsky 2006). The long-range cross-correlated processes could be defined this way as well. However, it is easy to show that the non-summability of cross-correlations easily arises from the definition of the cross-persistent processes through a power-law decay of the cross-correlation function.

Proposition 2.1 (Non-summability of cross-correlations). *Let's have two jointly wide-sense stationary processes $\{x_t\}$ and $\{y_t\}$ which are cross-persistent according to Definition 2.4. Then the cross-correlations are absolutely non-summable, i.e. $\sum_{j=-\infty}^{+\infty} |\rho_{xy}(j)| = +\infty$.*

Proof. We can write

$$\sum_{j=-\infty}^{+\infty} \rho_{xy}(j) = \rho_{xy}(0) + \sum_{j=1}^{+\infty} \rho_{xy}(j) + \sum_{j=-\infty}^{-1} \rho_{xy}(j). \tag{2.15}$$

Obviously, $\rho_{xy}(0)$ is finite and $-1 \leq \rho_{xy}(0) \leq 1$. We then have several scenarios.

First, the cross-correlations are positive and their decay is symmetric so that according to Definition 2.4, we have $\sum_{j=1}^{+\infty} \rho_{xy}(j) + \sum_{j=-\infty}^{-1} \rho_{xy}(j) \propto \sum_{j=1}^{+\infty} j^{-\gamma_{xy}}$ so that we can write

$$\sum_{j=-\infty}^{+\infty} |\rho_{xy}(j)| = \sum_{j=-\infty}^{+\infty} \rho_{xy}(j) \propto \sum_{j=1}^{+\infty} j^{-\gamma_{xy}}. \tag{2.16}$$

For monotonically decreasing positive function $j^{-\gamma_{xy}}$, the sum $\sum_{j=1}^{+\infty} j^{-\gamma_{xy}}$ converges if and only if $\int_1^{+\infty} j^{-\gamma_{xy}} dj$ is finite (MacLaurin 1742; Cauchy 1889). For the cross-persistent series with $0 < \gamma_{xy} < 1$, the integral diverges so that $\sum_{j=-\infty}^{+\infty} |\rho_{xy}(j)| = +\infty$. Note that according to the MacLaurin-Cauchy integral test, the sum diverges for all positive starting values. Therefore, it does not matter when the power-law decay starts as long as it is at some finite position.

Second, the cross-correlations are negative and their decay is symmetric. In the same manner as for the previous case, we have

$$\sum_{j=-\infty}^{+\infty} |\rho(j)| = \sum_{j=-\infty}^{+\infty} -\rho(j) = -\sum_{j=-\infty}^{+\infty} \rho(j) \tag{2.17}$$

and the sum diverges in the same manner as for the previous case.

Third, the cross-correlations are asymmetric. In any case, for the processes to be long-range cross-correlated, the cross-correlation function must be following a power law decay for at least one of its parts. However, as it is obvious from the previous two cases, the divergence of the sum of the whole cross-correlation function follows from the divergence of at least one of its halves. Therefore, the sum of the absolute values of the cross-correlations holds also for the asymmetric case. The absolute value of cross-correlations in the sum ensures that the correlations are not erased for an asymmetric power-law decay with different signs of correlations for two parts of the cross-correlation function. □

Note that the third part of the previous proof allows for various forms of the cross-correlation functions combining the power-law scaling, short-range cross-correlations or even uncorrelatedness for one half of the cross-correlation

function. The previous proposition thus covers various forms of long-range cross-correlated processes and we use this property in Chapter 4 to develop a test for a presence of long-range cross-correlations. A crucial construction point of the test is a different behavior of sums of the cross-correlation function for long-range and short-range cross-correlated processes. For the former, we have already shown that the sum of the absolute values of all cross-correlations diverges. But for the latter, the sum converges to a finite constant which is supported by the following proposition.

Proposition 2.2 (Summability of cross-correlations). *Let's have two jointly wide-sense stationary processes* $\{x_t\}$ *and* $\{y_t\}$ *which are short-range cross-correlated according to Definition 2.3. Then the cross-correlations are absolutely summable, i.e.* $\sum_{j=-\infty}^{+\infty} |\rho_{xy}(j)|$ *converges.*

Proof. In the similar steps as in the previous proof, we can write

$$\sum_{j=-\infty}^{+\infty} \rho_{xy}(j) = \rho_{xy}(0) + \sum_{j=1}^{+\infty} \rho_{xy}(j) + \sum_{j=-\infty}^{-1} \rho_{xy}(j). \tag{2.18}$$

and we have $-1 \leq \rho_{xy}(0) \leq 1$. We focus on a case when the cross-correlations are positive and symmetric. The other cases can be shown to hold as well in the manner of the previous proof.

As the cross-correlation function is symmetric, the processes are short-range cross-correlated according to Definition 2.3 and the cross-correlations are positive, we have

$$\sum_{j=-\infty}^{+\infty} |\rho_{xy}(j)| = \sum_{j=-\infty}^{+\infty} \rho_{xy}(j) \propto \sum_{j=1}^{+\infty} \exp(-j/\delta). \tag{2.19}$$

The last term is evidently a geometric progression with a common ratio $e^{-1/\delta}$. From Definition 2.3, we have $0 \leq \delta < +\infty$ so that $0 \leq \exp(-1/\delta) < 1$ and the geometric progression converges. We then arrive at

$$\sum_{j=1}^{+\infty} \exp(-j/\delta) = \frac{1}{1 - \exp(-1/\delta)} < +\infty \tag{2.20}$$

which together with Equation 2.19 concludes the proof. □

In Chapter 4, this distinction between short-range and long-range cross-correlated processes is used for a construction of the aggregate cross-correlations

test. In the next two sections, we focus on implications of the definition of the long-range cross-correlations on properties of partial sums and cross-power spectrum of the processes.

2.2.2 Scaling of covariances of the partial sums

For practical purposes, the analysis of the asymptotic behavior of cross-correlation function is rather complicated. It turns out that it is usually more convenient to study the behavior of partial sums of the processes and their cross-power spectrum. Let us first turn to the partial sums.

Definition 2.5 (Partial sum). Let's have a stationary process $\{x_t\}$ with $\langle x_t \rangle = 0$ and $\langle x_t^2 \rangle = \sigma_x^2 < +\infty$. Partial sum process $\{X_t\}$ is defined as

$$X_t = x_1 + x_2 + \ldots + x_t = \sum_{i=1}^{t} x_i. \tag{2.21}$$

Historically, long-range dependence was analyzed by Hurst (1951) using the rescaled range analysis (Mandelbrot & Wallis 1968), which is based on the assumption that the adjusted rescaled ranges of the partial sums of a zero mean process scale according to a power law. Other measures of variation have been used alongside the adjusted ranges to study long-term dependence, the most popular being the detrended fluctuation analysis (Peng *et al.* 1993; 1994; Kantelhardt *et al.* 2002) and various methods covered by Taqqu *et al.* (1995). We follow this logic for the long-range cross-correlated processes in the next proposition.

Proposition 2.3 (Partial sum covariance scaling). *Let's have stationary zero mean processes $\{x_t\}$ and $\{y_t\}$ with respective finite variances σ_x^2 and σ_y^2, and partial sums $\{X_t\}$ and $\{Y_t\}$ according to Definition 2.5. If processes $\{x_t\}$ and $\{y_t\}$ are long-range cross-correlated, the covariance between their partial sums scales as*

$$Cov(X_n, Y_n) \propto n^{2H_{xy}} \text{ as } n \to +\infty \tag{2.22}$$

where H_{xy} is the bivariate Hurst exponent. Moreover, it holds that $H_{xy} = 1 - \frac{\gamma_{xy}}{2}$.

Proof. Using the zero mean and stationarity properties of processes $\{x_t\}$ and

$\{y_t\}$, we can write the covariance of the partial sums as

$$\mathrm{Cov}(X_n, Y_n) = \langle X_n Y_n \rangle = \sigma_x \sigma_y \Big(n\rho_{xy}(0) + \sum_{k=1}^{n-1} (n-k)(\rho_{xy}(k) + \rho_{xy}(-k)) \Big)$$
$$\propto n\rho_{xy}(0) + \sum_{k=1}^{n-1} (n-k)(\rho_{xy}(k) + \rho_{xy}(-k)). \quad (2.23)$$

Now, assuming that $\rho_{xy}(k)$ is symmetric for $k > 0$ and $k < 0$, we have

$$\mathrm{Cov}(X_n, Y_n) \propto n\rho_{xy}(0) + n \sum_{k=1}^{n-1} \rho_{xy}(k) - \sum_{k=1}^{n-1} k\rho_{xy}(k). \quad (2.24)$$

Using Definition 2.4 and approximating the infinite sums with definite integrals according to the Euler–MacLaurin integration formula (Euler 1738; MacLaurin 1742), we get

$$n \sum_{k=1}^{n-1} \rho_{xy}(k) \propto n \sum_{k=1}^{n-1} k^{-\gamma_{xy}} \approx n \int_1^n k^{-\gamma_{xy}} dk \propto n^{2-\gamma_{xy}}, \quad (2.25)$$

$$\sum_{k=1}^{n-1} k\rho_{xy}(k) \propto \sum_{k=1}^{n-1} k^{1-\gamma_{xy}} \approx \int_1^n k^{1-\gamma_{xy}} dk \propto n^{2-\gamma_{xy}}. \quad (2.26)$$

Finally, we use that the linear growth of $n\rho_{xy}(0)$ is asymptotically dominated by the power-law growth in the latter terms, i.e. using the l'Hôpital's rule we have

$$\lim_{n \to +\infty} \frac{n^{2-\gamma_{xy}}}{n\rho_{xy}(0)} = \lim_{n \to +\infty} \frac{(2-\gamma_{xy}) n^{1-\gamma_{xy}}}{\rho_{xy}(0)} = +\infty \text{ for } 0 < \gamma_{xy} < 1 \quad (2.27)$$

and we get

$$\mathrm{Cov}(X_n, Y_n) \propto n^{2-\gamma_{xy}} \text{ as } n \to +\infty. \quad (2.28)$$

Note that the substitutions in Equation 2.25 and Equation 2.26 from $\sum_{k=1}^{n-1} \rho_{xy}(k)$ to $\sum_{k=1}^{n-1} k^{-\gamma_{xy}}$ are done for k between 1 and $n-1$ without a loss on generality as we are interested in the asymptotic properties of $\mathrm{Cov}(X_n, Y_n)$.

From Equation 2.22 and 2.28, we have $2H_{xy} = 2 - \gamma_{xy}$ so that

$$H_{xy} = 1 - \frac{\gamma_{xy}}{2}. \quad (2.29)$$

For the asymmetric cross-correlation function, the results do not differ sig-

nificantly. Instead of Equation 2.24, we have

$$\mathrm{Cov}(X_n, Y_n) \approx n\rho_{xy}(0) + \underbrace{n\sum_{k=1}^{n-1} k^{-\gamma^1_{xy}} - \sum_{k=1}^{n-1} k^{-\gamma^1_{xy}+1}}_{\propto n^{2-\gamma^1_{xy}}} + \underbrace{n\sum_{k=1}^{n-1} k^{-\gamma^2_{xy}} - \sum_{k=1}^{n-1} k^{-\gamma^2_{xy}+1}}_{\propto n^{2-\gamma^2_{xy}}}, \tag{2.30}$$

where the approximate proportionality comes from Equation 2.25 and 2.26. Asymptotically, the power-law scaling is dominated by the higher exponent, i.e. the lower γ_{xy}. For $\gamma^1_{xy} < \gamma^2_{xy}$, we have $\mathrm{Cov}(X_n, Y_n) \sim n^{2-\gamma^1_{xy}}$ and vice versa. Note that the lower γ_{xy} is connected to the higher bivariate Hurst exponent H_{xy} which implies that the scaling of covariances is dominated by the stronger cross-persistence. □

Non-summability of cross-correlations in Proposition 2.1 and divergence of the covariance of the partial sums in Proposition 2.3 leads to the following proposition.

Proposition 2.4 (Diverging limit of covariance of partial sums). *For two jointly stationary long-range cross-correlated processes according to Definition 2.4, $\{x_t\}$ and $\{y_t\}$ and their respective partial sums $\{X_t\}$ and $\{Y_t\}$, it holds that*

$$\lim_{n\to+\infty} \frac{Cov(X_n, Y_n)}{n} = +\infty. \tag{2.31}$$

Proof. From Proposition 2.3, we have

$$\lim_{n\to+\infty} \frac{\mathrm{Cov}(X_n, Y_n)}{n} \propto \lim_{n\to+\infty} \frac{n^{2H_{xy}}}{n} = \lim_{n\to+\infty} \frac{n^{2-\gamma_{xy}}}{n} = \lim_{n\to+\infty} n^{1-\gamma_{xy}} = +\infty \text{ for } 0 < \gamma_{xy} < 1. \tag{2.32}$$

□

Divergence in Equation 2.31 is parallel to the divergence of the variance of the partial sums for the long-range dependent processes as shown by Samorodnitsky (2006) and can thus be seen as a sign of long-range cross-correlations. We use this proposition to construct a test for long-range cross-correlated processes in Chapter 4. However, distinguishing between the short- and long-range cross-correlated processes only makes sense if the diverging limit in Proposition 2.31 is not the case for the short-range cross-correlated processes. The following proposition and its proof indeed show so.

Proposition 2.5 (Converging limit of covariance of partial sums). *For two jointly stationary short-range cross-correlated processes according to Definition 2.3, $\{x_t\}$ and $\{y_t\}$, and their respective partial sums $\{X_t\}$ and $\{Y_t\}$, the expression*

$$\lim_{n\to+\infty} \frac{Cov(X_n, Y_n)}{n} \tag{2.33}$$

converges.

Proof. In accordance with the proof of Proposition 2.3, we assume a symmetric cross-correlation function so that we can write

$$\mathrm{Cov}(X_n, Y_n) \propto n\rho_{xy}(0) + n\sum_{k=1}^{n-1} \rho_{xy}(k) - \sum_{k=1}^{n-1} k\rho_{xy}(k). \tag{2.34}$$

For an asymmetric case, the proof is parallel. It holds that

$$\lim_{n\to+\infty} \frac{\mathrm{Cov}(X_n, Y_n)}{n} \propto \lim_{n\to+\infty} \left(\rho_{xy}(0) + \sum_{k=1}^{n-1} \rho_{xy}(k) - \frac{1}{n}\sum_{k=1}^{n-1} k\rho_{xy}(k) \right). \tag{2.35}$$

Solving the sums separately with a use of short-range cross-correlations definition (Definition 2.3), we get

$$\sum_{k=1}^{n-1} \rho_{xy}(k) \propto \sum_{k=1}^{n-1} \exp\left(-\frac{k}{\delta}\right) \propto \frac{1-\exp\left(-\frac{n}{\delta}\right)}{1-\exp\left(-\frac{1}{\delta}\right)} \tag{2.36}$$

$$\sum_{k=1}^{n-1} k\rho_{xy}(k) \propto \sum_{k=1}^{n-1} k\exp\left(-\frac{k}{\delta}\right) = \exp\left(-\frac{1}{\delta}\right) - n\exp\left(-\frac{n}{\delta}\right) + (n-1)\exp\left(-\frac{n+1}{\delta}\right). \tag{2.37}$$

Substituting back to Equation 2.35, we get

$$\lim_{n\to+\infty} \frac{\mathrm{Cov}(X_n, Y_n)}{n} \propto \lim_{n\to+\infty} \left[\rho_{xy}(0) + \frac{1-\exp\left(-\frac{n}{\delta}\right)}{1-\exp\left(-\frac{1}{\delta}\right)} - \frac{\exp\left(-\frac{1}{\delta}\right)}{n} + \frac{n}{n}\exp\left(-\frac{n}{\delta}\right) + \frac{n-1}{n}\exp\left(-\frac{n+1}{\delta}\right) \right] = \rho_{xy}(0) + \frac{1}{1-\exp\left(-\frac{1}{\delta}\right)} \tag{2.38}$$

and the limit evidently converges for $0 \leq \delta < +\infty$ which concludes the proof. □

The statistical properties of covariance of the partial sums of short- and long-range cross-correlated processes give enough information to construct estimators of the bivariate Hurst exponent H_{xy} via its asymptotic scaling, which is utilized in Chapter 5 in the detrended cross-correlation analysis and the height cross-correlation analysis and it can also be used to distinguish between the two types of processes through converge or divergence of the covariances.

Up to this point, we have focused on the time domain properties of long-range cross-correlated series. Practically unequivocally, the basic properties of the univariate long-term memory translate into the bivariate setting regardless the symmetry of the cross-correlation function. Implications of the cross-persistence in the frequency domain are examined in the following section.

2.2.3 Cross-power spectrum divergence

Long-range dependent processes are characterized by a divergent power spectrum for low frequencies, specifically characterized by a power-law scaling, which emphasizes that the dynamics of the process is dominated by the behavior at high scales. In the same manner, behavior of the series in the bivariate setting is dominated by the dynamics at low frequencies (high scales) so that the scaling of the cross-power spectrum should behave similarly. Without imposing any additional assumptions on the underlying processes, we present a proposition which describes a scaling of the magnitude of the cross-power spectrum.

Proposition 2.6 (Cross-power spectrum of LRCC processes). *For two jointly stationary long-range cross-correlated processes according to Definition 2.5,* $\{x_t\}$ *and* $\{y_t\}$*, the cross-power spectrum diverges at the origin* $\lambda \to 0+$ *and the divergence is of a form*

$$|f_{xy}(\lambda)| \propto \lambda^{-\beta_{xy}}, \tag{2.39}$$

where β_{xy} *is a scaling exponent. Moreover, it holds that* $\beta_{xy} = 1 - \gamma_{xy}$*.*

Proof. We can decompose the cross-power spectrum as

$$f_{xy}(\lambda) \propto \sum_{n=-\infty}^{\infty} \rho_{xy}(n) \exp(-i\lambda n) = \rho_{xy}(0) + \sum_{n=1}^{\infty} \rho_{xy}(n) \exp(-i\lambda n) + \sum_{n=1}^{\infty} \rho_{xy}(-n) \exp(i\lambda n). \tag{2.40}$$

Let us first start with a case of the symmetric cross-correlation function, i.e. $\rho_{xy}(n) = \rho_{xy}(-n)$. From Equation 2.40 and using that $\exp(iz) = \cos(z) + i\sin(z)$, we can write

$$f_{xy}(\lambda) \propto \sum_{n=1}^{\infty} \rho_{xy}(n)[\exp(i\lambda n) + \exp(-i\lambda n)] = \sum_{n=1}^{\infty} \rho_{xy}(n)[\cos(\lambda n) + \cos(-\lambda n) + i\sin(\lambda n) + i\sin(-\lambda n)] = 2\sum_{n=1}^{\infty} \rho_{xy}(n)\cos(\lambda n) \quad (2.41)$$

Using Definition 2.4 and approximating the infinite sum with the definite integral, we get

$$f_{xy}(\lambda) \propto \sum_{n=1}^{\infty} n^{-\gamma_{xy}}\cos(\lambda n) \approx \int_1^{\infty} n^{-\gamma_{xy}}\cos(\lambda n)dn = \frac{F_p^q\left(\frac{1-\gamma_{xy}}{2}, \left[\frac{1}{2}, \frac{3-\gamma_{xy}}{2}\right], -\frac{\lambda^2}{4}\right)}{-1+\gamma_{xy}} + \frac{\Gamma(1-\gamma_{xy})\sin\left(\frac{\gamma_{xy}\pi}{2}\right)}{|\lambda|^{1-\gamma_{xy}}}, \quad (2.42)$$

where $F_p^q(\bullet, \bullet, \bullet)$ is the generalized hypergeometric function. We can use that

$$\lim_{\lambda \to 0+} F_p^q\left(\frac{1-\gamma_{xy}}{2}, \left[\frac{1}{2}, \frac{3-\gamma_{xy}}{2}\right], -\frac{\lambda^2}{4}\right) = 1 \text{ for } 0 < \gamma_{xy} < 1, \quad (2.43)$$

and that $\Gamma(1-\gamma_{xy})\sin\left(\frac{\gamma_{xy}\pi}{2}\right)$ is a constant for $0 < \gamma_{xy} < 1$ to argue that close to the origin from the right, we have $f_{xy}(\lambda) \propto \lambda^{-(1-\gamma_{xy})}$. Evidently from Equation 2.42, we have $f_{xy}(\lambda) \in \mathbb{R}^+$ so that $f_{xy}(\lambda) = |f_{xy}(\lambda)|$ and thus

$$|f_{xy}(\lambda)| \propto \lambda^{-(1-\gamma_{xy})}. \quad (2.44)$$

From Equation 2.39 and Equation 2.44, we have $\beta_{xy} = 1 - \gamma_{xy}$. For $0 < \gamma_{xy} < 1$, we have

$$\lim_{\lambda \to 0+} |f_{xy}(\lambda)| \propto \lim_{\lambda \to 0+} \lambda^{-(1-\gamma_{xy})} = +\infty. \quad (2.45)$$

so that the cross-power spectrum diverges at the origin, which concludes the proof for the symmetric case.

When the cross-correlation function $\rho_{xy}(n)$ is asymmetric, the cross-power spectrum is complex-valued, i.e. the sine functions are not erased in Equa-

tion 2.41 and the proof is more complicated. Starting with the cross-power spectrum written as

$$f_{xy}(\lambda) \propto \sum_{n=1}^{\infty} \rho_{xy}(n) \exp(-i\lambda n) + \sum_{n=1}^{\infty} \rho_{xy}(-n) \exp(i\lambda n), \tag{2.46}$$

we need to solve the sums separately. Assuming that $\rho_{xy}(n) \propto n^{-\gamma_{xy}^{+}}$ for $n > 0$ and $\rho_{xy}(n) \propto n^{-\gamma_{xy}^{-}}$ for $n < 0$, we can write the first sum in Equation 2.46 as

$$\sum_{n=1}^{\infty} \rho_{xy}(n) \exp(-i\lambda n) \propto \sum_{n=1}^{\infty} n^{-\gamma_{xy}^{+}} [\cos(-\lambda n) + i \sin(-\lambda n)], \tag{2.47}$$

and the second sum in Equation 2.46 as

$$\sum_{n=1}^{\infty} \rho_{xy}(-n) \exp(i\lambda n) \propto \sum_{n=1}^{\infty} n^{-\gamma_{xy}^{-}} [\cos(\lambda n) + i \sin(\lambda n)]. \tag{2.48}$$

By putting Equation 2.47 and 2.48 together, we obtain

$$\begin{aligned} f_{xy}(\lambda) \propto \sum_{n=1}^{\infty} \Big(\cos(\lambda n) \big[n^{-\gamma_{xy}^{+}} + n^{-\gamma_{xy}^{-}} \big] + i \big[n^{-\gamma_{xy}^{+}} \sin(-\lambda n) + n^{-\gamma_{xy}^{-} \sin(\lambda n)} \big] \Big) = \\ \sum_{n=1}^{\infty} \big(\cos(\lambda n) \big[n^{-\gamma_{xy}^{+}} + n^{-\gamma_{xy}^{-}} \big] \big) + i \sum_{n=1}^{\infty} \big[n^{-\gamma_{xy}^{+}} \sin(-\lambda n) + n^{-\gamma_{xy}^{-}} \sin(\lambda n) \big]. \end{aligned} \tag{2.49}$$

The infinite sums are solved separately using the approximation by the definite integrals for $0 < \lambda \leq \pi$ and $0 < \gamma_{xy}^{+}, \gamma_{xy}^{-} < 1$:

$$\begin{aligned} \sum_{n=1}^{\infty} \big(\cos(\lambda n) \big[n^{-\gamma_{xy}^{+}} + n^{-\gamma_{xy}^{-}} \big] \big) \approx \int_{1}^{\infty} \cos(\lambda n) \big[n^{-\gamma_{xy}^{+}} + n^{-\gamma_{xy}^{-}} \big] dn = \\ \frac{F_p^q \left(\frac{1-\gamma_{xy}^{+}}{2}, \left[\frac{1}{2}, \frac{3-\gamma_{xy}^{+}}{2} \right], -\frac{\lambda^2}{4} \right)}{\gamma_{xy}^{+} - 1} + \frac{F_p^q \left(\frac{1-\gamma_{xy}^{-}}{2}, \left[\frac{1}{2}, \frac{3-\gamma_{xy}^{-}}{2} \right], -\frac{\lambda^2}{4} \right)}{\gamma_{xy}^{-} - 1} + \\ \frac{\Gamma(1-\gamma_{xy}^{+}) \sin\left(\frac{\gamma_{xy}^{+} \pi}{2} \right)}{\lambda^{1-\gamma_{xy}^{+}}} + \frac{\Gamma(1-\gamma_{xy}^{-}) \sin\left(\frac{\gamma_{xy}^{-} \pi}{2} \right)}{\lambda^{1-\gamma_{xy}^{-}}} \end{aligned} \tag{2.50}$$

$$\sum_{n=1}^{\infty}\left[n^{-\gamma_{xy}^{+}}\sin(-\lambda n)+n^{-\gamma_{xy}^{-}}\sin(\lambda n)\right]\approx \int_{1}^{\infty}\left[n^{-\gamma_{xy}^{+}}\sin(-\lambda n)+n^{-\gamma_{xy}^{-}}\sin(\lambda n)\right]dn = -\frac{\lambda F_p^q\left(\frac{2-\gamma_{xy}^{+}}{2},\left[\frac{3}{2},\frac{4-\gamma_{xy}^{+}}{2}\right],-\frac{\lambda^2}{4}\right)}{\gamma_{xy}^{+}-2}+\frac{\lambda F_p^q\left(\frac{2-\gamma_{xy}^{-}}{2},\left[\frac{3}{2},\frac{4-\gamma_{xy}^{-}}{2}\right],-\frac{\lambda^2}{4}\right)}{\gamma_{xy}^{-}-2}- \frac{\Gamma(1-\gamma_{xy}^{+})\cos\left(\frac{\gamma_{xy}^{+}\pi}{2}\right)}{\lambda^{1-\gamma_{xy}^{+}}}+\frac{\Gamma(1-\gamma_{xy}^{-})\cos\left(\frac{\gamma_{xy}^{-}\pi}{2}\right)}{\lambda^{1-\gamma_{xy}^{-}}} \quad (2.51)$$

As λ approaches zero from the right, all the generalized hypergeometric functions in Equation 2.50 and 2.51 approach a unity. Therefore, we are left with two scaling relationships – $\lambda^{-(1-\gamma_{xy}^{+})}$ and $\lambda^{-(1-\gamma_{xy}^{-})}$ – for both Equation 2.50 and 2.51. Assuming that $\gamma_{xy}^{+}<\gamma_{xy}^{-}$, this implies that as λ approaches the origin, the scaling with the exponent γ_{xy}^{+} dominates the scaling with the exponent γ_{xy}^{-} because

$$\lim_{\lambda\to 0+}\frac{\lambda^{\gamma_{xy}^{+}-1}}{\lambda^{\gamma_{xy}^{-}-1}}=\lim_{\lambda\to 0+}\lambda^{\gamma_{xy}^{+}-\gamma_{xy}^{-}}=+\infty \text{ as } \gamma_{xy}^{+}<\gamma_{xy}^{-}. \quad (2.52)$$

Therefore, as $\lambda\to 0+$, we can write

$$\sum_{n=1}^{\infty}\cos(\lambda n)\left[n^{-\gamma_{xy}^{+}}+n^{-\gamma_{xy}^{-}}\right]\propto\lambda^{\gamma_{xy}^{+}-1} \quad (2.53)$$

$$\sum_{n=1}^{\infty}\left[n^{-\gamma_{xy}^{+}}\sin(-\lambda n)+n^{-\gamma_{xy}^{-}}\sin(\lambda n)\right]\propto\lambda^{\gamma_{xy}^{+}-1} \quad (2.54)$$

so that from Equation 2.49, both real and imaginary part of the cross-power spectrum follow the power law with the same scaling exponent, i.e.

$$Re[f_{xy}(\lambda)]\propto\lambda^{\gamma_{xy}^{+}-1} \quad (2.55)$$

$$Im[f_{xy}(\lambda)]\propto\lambda^{\gamma_{xy}^{+}-1}. \quad (2.56)$$

Using that for the complex numbers $|a+ib|=\sqrt{a^2+b^2}$ and labeling $\lambda_{xy}\equiv\lambda_{xy}^{+}$, we have

$$|f_{xy}(\lambda)|\propto\lambda^{\gamma_{xy}-1}=\lambda^{-(1-\gamma_{xy})}, \quad (2.57)$$

which implies the same divergence and the relationship between γ_{xy} and β_{xy} as in Equation 2.45. This concludes the proof also for the asymmetric case. Note that apart from the case with two different scaling exponents γ_{xy}^{+} and γ_{xy}^{-}, we can have various combinations of memories in the cross-correlation function.

However, the correlations influence only high frequencies of the spectrum apart from the case of the long-range cross-correlations. □

The scaling law is strongly connected to the definition of long-range cross-correlated processes via the power-law decay of cross-correlation function. We stress here that the cross-power spectrum magnitude scaling is implied from the scaling of the cross-correlation function of the cross-persistent processes and the reversed implication is not so obvious without further assumptions which is mainly due to potentially asymmetric cross-power spectrum. This is a main difference from the univariate case with by definition symmetric power spectrum.

The definition of the cross-persistence through the asymptotic scaling of the cross-correlation function implies two additional scaling laws which are also partially characteristic for the univariate long-term memory – power spectrum divergence (cross-power spectrum magnitude divergence) and scaling of variance (covariance) of partial sums. A triangle relationship among the scaling exponent of the power-law decay of the cross-correlation function, γ_{xy}, the scaling exponent of the covariance of the partial sums, the bivariate Hurst exponent H_{xy}, and the scaling exponent of the diverging at the origin magnitude of the cross-power spectrum, β_{xy}, is also implied and is summarized in the following proposition.

Proposition 2.7 (Scaling triangle). *For two jointly stationary long-range cross-correlated processes $\{x_t\}$ and $\{y_t\}$ with γ_{xy} according to Definition 2.5, with cross-power spectrum $f_{xy}(\lambda)$ and scaling exponent β_{xy} and with partial sums with the bivariate Hurst exponent H_{xy}, it holds that*

$$H_{xy} = 1 - \frac{\gamma_{xy}}{2}, \tag{2.58}$$

$$\beta_{xy} = 1 - \gamma_{xy}, \tag{2.59}$$

$$H_{xy} = \frac{1 + \beta_{xy}}{2}. \tag{2.60}$$

Proof. The first two points are parts of Proposition 2.3 and 2.6. The third point easily arises from the other two:

$$H_{xy} = 1 - \frac{\gamma_{xy}}{2} = 1 - \frac{1 - \beta_{xy}}{2} = \frac{1 + \beta_{xy}}{2} \tag{2.61}$$

□

The relationship among the different scaling measures associated with the long-range cross-correlations leads to a crucial proposition about existence of long-range cross-correlated processes. To proof the propositions, we need to define a squared spectrum coherence.

Definition 2.6 (Squared spectrum coherence). Let $\{x_t\}$ and $\{y_t\}$ be jointly stationary processes with existing $f_{xy}(\lambda)$, $f_{xx}(\lambda)$ and $f_{yy}(\lambda)$ at frequency $0 \leq \lambda < \pi$. Squared spectrum coherency is defined as

$$K_{xy}^2(\lambda) = \frac{|f_{xy}(\lambda)|^2}{f_{xx}(\lambda) f_{yy}(\lambda)}. \tag{2.62}$$

The squared coherence can be understood as a squared correlation between processes $\{x_t\}$ and $\{y_t\}$ at frequency λ. Returning to the existence of long-range cross-correlated processes, we split the propositions into two cases to make the implications more evident – for long-range correlated processes and for short-range correlated processes (including the serially uncorrelated processes).

Proposition 2.8 (Bivariate Hurst exponent restriction I). *Let $\{x_t\}$ and $\{y_t\}$ be jointly stationary long-range cross-correlated processes with $H_x, H_y > 0.5$. It holds that $H_{xy} \leq \frac{H_x + H_y}{2}$.*

Proof. Let's proof the proposition by contradiction. Recall that for the squared spectrum coherence $K_{xy}^2(\lambda)$, it holds that $0 \leq K_{xy}^2(\lambda) \leq 1$ for all λ (Wei 2006). Assume that $H_{xy} > \frac{H_x + H_y}{2}$. The long-range cross-correlation property given in Proposition 2.6 and the relationship to the bivariate Hurst exponent in Proposition 2.7 give $|f_{xy}(\lambda)| \propto \lambda^{1-2H_{xy}}$, $f_{xx}(\lambda) \propto \lambda^{1-2H_x}$ and $f_{yy}(\lambda) \propto \lambda^{1-2H_y}$ near the origin. The coherence is then given as

$$K_{xy}^2(\lambda) \propto \frac{\lambda^{2(1-2H_{xy})}}{\lambda^{1-2H_x}\lambda^{1-2H_y}} = \lambda^{2(H_x+H_y-2H_{xy})}. \tag{2.63}$$

Close to the origin, we have

$$\lim_{\lambda \to 0+} K_{xy}^2(\lambda) \propto \lim_{\lambda \to 0+} \lambda^{2(H_x+H_y-2H_{xy})} = +\infty \tag{2.64}$$

for $H_{xy} > \frac{H_x + H_y}{2}$ which is a contradiction as $0 \leq K_{xy}^2(\lambda) \leq 1$ for $0 \leq \lambda \leq \pi$. □

Proposition 2.9 (Bivariate Hurst exponent restriction II). *Let $\{x_t\}$ and $\{y_t\}$ be jointly stationary processes with finite at origin respective spectra $f_{xx}(\lambda)$ and $f_{yy}(\lambda)$. Then the processes $\{x_t\}$ and $\{y_t\}$ are not long-range cross-correlated.*

Proof. In the same manner as in the previous proof, we use a contradiction. Assume that processes $\{x_t\}$ and $\{y_t\}$ are long-range cross-correlated, i.e. $H_{xy} > 0.5$. This implies that close to the origin, the coherence is

$$\lim_{\lambda \to 0+} K_{xy}^2(\lambda) \propto \lim_{\lambda \to 0+} \lambda^{2(1-2H_{xy})} = +\infty \tag{2.65}$$

for $H_{xy} > 0.5$ which again contradicts $0 \leq K_{xy}^2(\lambda) \leq 1$ for $0 \leq \lambda \leq \pi$. □

These two propositions have serious consequences. First, if two processes are serially uncorrelated (or only short-range cross-correlated), they cannot be cross-persistent. Second, if the separate processes are long-range correlated, they can only be cross-correlated up to the point where the bivariate Hurst exponent is the average of the separate univariate Hurst exponents. Third, the cross-correlated processes do not necessarily possess the bivariate Hurst exponent equal to the average of the separate Hurst exponents, which has been identified repeatedly in the literature (Podobnik & Stanley 2008; Kristoufek 2011). Fourth, if the bivariate Hurst exponent is found to be higher than the average of the two separate Hurst exponents (He & Chen 2011a;b), the results are spurious, probably due to relatively poor finite sample properties of the frequently used estimators as we show in Chapter 5. The implications of Proposition 2.8 are well illustrated in Section 3.3 where we introduce a model for which we can alter H_x, H_y and H_{xy} as long as $H_{xy} \leq \frac{H_x+H_y}{2}$.

Proofs to Proposition 2.8 and 2.9 uncover an interesting possibility, which has been recently discussed by Sela & Hurvich (2012) – the power law coherency. This situation occurs when $0.5 < H_{xy} < \frac{H_x+H_y}{2}$ and as is evident from the previous two proofs, the squared spectrum coherence follows a power law and goes to zero at $\lambda \to 0+$. Sela & Hurvich (2012) propose a new model with the power law coherency and they call it an anti-cointegration as the separate processes are long-range correlated – fractionally integrated – but are pairwise uncorrelated in a long-term horizon (at low frequencies). This is in evident opposition with (fractional) cointegration for which it holds that $K_{xy}^2(\lambda) = 1$ as $\lambda \to 0+$. To measure the strength of the power law coherency, the authors propose to use $d_\rho = d_{12} - \frac{d_1+d_2}{2}$ where d_{12}, d_1 and d_2 are fractional integration parameters for the joint long-term memory and the separate long-term memories, respectively. As we mainly function with the Hurst exponent definitions, we will use the measure $H_\rho = H_{xy} - \frac{H_x+H_y}{2} = d_{12} + \frac{d_1+d_2}{2} = d_\rho$ so that these are equivalent. If it holds that $H_{xy} = \frac{H_x+H_y}{2}$, we have $H_\rho = 0$.

To summarize the current chapter, we have shown that the basic properties

of long-range correlated processes also hold for the bivariate setting (apart from the last two propositions). This is also an important finding as it enables us to construct tests and estimators for the bivariate setting which are motivated and parallel to the ones for the univariate setting. Before turning to the tests for the long-range cross-correlated processes in Chapter 4 and the estimators of the bivariate Hurst exponent in Chapter 5, we focus on several long-range cross-correlated processes which are later used for the Monte Carlo simulations.

Chapter 3

Analytic results for specific processes

Existence of statistically significant long-range cross-correlations between various series is a fascinating phenomenon important for modeling and forecasting time series. Several processes that possess such long-term correlations have been proposed in the literature. The most frequently discussed and applied ones are the multivariate generalizations of the well-established fractionally integrated ARMA processes (usually labeled as FARIMA and ARFIMA) – VARFIMA or MVARFIMA processes, compare Ravishanker & Ray (1997), Martin & Wilkins (1999), Nielsen (2004a), Shimotsu (2007), Tsay (2010), and Nielsen (2011) – and fractional Gaussian noise processes or fractional Brownian motions, which are their integrated version (these are labeled as fGn and fBm in the literature, respectively). The construction of the multivariate ARFIMA process implies that the bivariate Hurst exponent is the average of the separate Hurst exponents (Nielsen 2011). The same property holds for the fractional Brownian motion (Amblard & Coeurjolly 2011). The long-range cross-correlations thus simply arise from the specification of these processes.

Lobato (1997) and then in some detail Sela & Hurvich (2009) discuss two types of fractionally integrated models – VARFI and FIVAR. VARFI is a vector autoregressive model with fractionally integrated innovations (or error terms). The pairwise processes are then long-range correlated but only short-range cross-correlated so that the bivariate Hurst exponent is equal to 0.5 and is thus lower than the average of the separate Hurst exponents. Reversely, FIVAR consists of fractionally integrated processes with innovations that come from a VAR model. The pairwise processes are then both long-range correlated and

long-range cross-correlated with the bivariate Hurst exponent being equal to the average of the separate Hurst exponents as for the previous cases. Sela & Hurvich (2012) propose an anti-cointegration model, which is in fact a linear combination of ARFIMA processes with a subset of innovations (but not all pairs) being identical across the two processes. The model allows to control the separate Hurst exponents as well as the bivariate Hurst exponent as long as it is lower or equal to the average of the separate parameters and it allows for the power law coherency.

Nielsen (2004b) and Sela (2010) discuss the case of the fractional cointegration, i.e. the case when both processes are fractionally integrated of the same order and there exists a linear combination of them which is stationary and fractionally integrated of a lower order, in the long-term memory setting. Both Nielsen (2004b) and Sela (2010) show that the coherence of the processes is equal to unity which implies that the bivariate Hurst exponent is the same as the separate Hurst exponents and so again is equal to their average.

In this chapter, we show how the power-law scaling emerges from the definition of several processes. To keep the derivations instructive, we stick to the bivariate case. Also, the main focus is put on utilizing ARFIMA-based processes rather than the generalizations of the fractional Gaussian noise due to simplicity of ARFIMA construction and its straightforward spectrum (see Goddard & Onali (2012) for comparison and Lavancier *et al.* (2009), Amblard & Coeurjolly (2011), Coeurjolly *et al.* (2012), and Amblard *et al.* (2012) for details on the multivariate generalizations of fGn). We start with the bivariate ARFIMA process with a simple correlation structure of the innovations – ARFIMA processes with correlated innovations. We then focus on the case when one of the processes is ARFIMA and the other is a simple AR(1) process with innovations being correlated again. Last but not least, we introduce a new kind of a bivariate process which we call the mixed-correlated ARFIMA process. The process is in a sense a generalization of the anti-cointegration model of Sela & Hurvich (2012). The process allows to control for the separate and bivariate Hurst exponents as long as the bivariate one is not higher than the average of the separate ones, additionally allows for short-range dependence and encompasses the fractional cointegration as a special case as well.

Each type of processes is represented as a moving average of infinite order for uncomplicated computations. For each process, we discuss stationarity and joint stationarity so that they cope with our definition of long-range cross-correlations in Definition 2.4. A pattern of cross-correlations is presented and

used for a construction of cross-power spectrum. The relationship between the bivariate Hurst exponent and separate Hurst exponents is demonstrated on the scaling of cross-correlation function, which is obtained from the cross-power spectrum using the inverse Fourier transform.

3.1 ARFIMA(0,d,0) processes with correlated innovations

We start with a case of two ARFIMA(0,d,0) processes with correlated innovations. In general, the correlation structure of the innovations can have various specifications as long as the processes remain stationary (see e.g. Samorodnitsky (2006) for details) so that these can be also pairwise correlated, autocorrelated or cross-correlated. For sake of simplicity, we stick to the simplest specification leading to the long-range cross-correlated processes – pairwise correlated innovations.

In the case of ARFIMA processes with correlated innovations, we describe the procedure step-by-step and in detail. For the following types of processes, we use the same principles and thus only partial derivation is reported.

ARFIMA processes with correlated innovations are defined simply as two ARFIMA(0,d,0) processes with parameters d_1 and d_2 and specific correlation structure so that we have

$$x_t = \sum_{n=0}^{\infty} a_n(d_1)\varepsilon_{t-n} \tag{3.1}$$

$$y_t = \sum_{n=0}^{\infty} a_n(d_2)\nu_{t-n} \tag{3.2}$$

where

$$a_n(d) = \frac{\Gamma(|n|+d)}{\Gamma(|n|+1)\Gamma(d)} \tag{3.3}$$

and innovations are characterized by

$$\begin{aligned}\langle\varepsilon_t\rangle = \langle\nu_t\rangle &= 0\\ \langle\varepsilon_t^2\rangle &= \sigma_\varepsilon^2 < +\infty\\ \langle\nu_t^2\rangle &= \sigma_\nu^2 < +\infty\\ \langle\varepsilon_t\varepsilon_{t-n}\rangle = \langle\nu_t\nu_{t-n}\rangle = \langle\varepsilon_t\nu_{t-n}\rangle &= 0 \text{ for } n \neq 0\\ \langle\varepsilon_t\nu_t\rangle &= \sigma_{\varepsilon\nu} < +\infty. \end{aligned} \tag{3.4}$$

Note that we use a slightly different definition of $a_n(d)$ which is more flexible for cross-correlations analysis and covers the standard definition without the absolute values on the right-hand side. The function is thus symmetric, $a_n(d) = a_{-n}(d)$.

We thus start with a quite uncomplicated dependence between series – separate series $\{x_t\}$ and $\{y_t\}$ are long-range dependent and their innovations are serially uncorrelated but pairwise correlated.

Before turning to the cross-correlations between processes $\{x_t\}$ and $\{y_t\}$, we shortly focus on means and variances of the processes, which obviously enter all cross-correlations, to make the calculations are more transparent. The mean values of the processes are evidently equal to zero as they are a linear combination of innovations with zero mean so that

$$\langle x_t\rangle = \langle y_t\rangle = 0. \tag{3.5}$$

Variance can be then taken as an expected value of the squared process leading to

$$\begin{aligned}\mathrm{var}(x_t) = \langle x_t^2\rangle = \langle(a_0(d_1)\varepsilon_t + a_1(d_1)\varepsilon_{t-1} + \ldots)(a_0(d_1)\varepsilon_t + a_1(d_1)\varepsilon_{t-1} + \ldots)\rangle =\\ \sum_{j=0}^{\infty} a_j^2(d_1)\langle\varepsilon_t^2\rangle = \sigma_\varepsilon^2\sum_{j=0}^{\infty} a_j^2(d_1) \propto \sigma_\varepsilon^2 < +\infty\end{aligned} \tag{3.6}$$

using that $\sum_{j=0}^{\infty} a_j^2(d)$ converges for $0 \leq d \leq 0.5$ (Samorodnitsky 2006, p. 45-46). In the same manner, we have

$$\mathrm{var}(y_t) = \langle y_t^2\rangle \propto \sigma_\nu^2 < +\infty. \tag{3.7}$$

For convenience, we write $\mathrm{var}(x_t) \equiv \sigma_x^2$ and $\mathrm{var}(y_t) \equiv \sigma_y^2$. Variance of both processes is thus independent of t and is finite. Sowell (1992) and Bertelli

& Caporin (2002) show that the auto-covariance function of ARFIMA(0,d,0) process is

$$\mathrm{Cov}(x_t, x_{t-n}) = \sigma_\varepsilon^2 \frac{\Gamma(n+d_1)\Gamma(1-2d_1)}{\Gamma(n+1-d_1)\Gamma(1-d_1)\Gamma(d_1)} \propto \sigma_\varepsilon^2 < +\infty \tag{3.8}$$

$$\mathrm{Cov}(y_t, y_{t-n}) = \sigma_\nu^2 \frac{\Gamma(n+d_2)\Gamma(1-2d_2)}{\Gamma(n+1-d_2)\Gamma(1-d_2)\Gamma(d_2)} \propto \sigma_\nu^2 < +\infty \tag{3.9}$$

The auto-covariances are thus also independent of t and finite so that the processes are wide-sense stationary. The structure of cross-correlations is independent of t, as in detail shown in Appendix A.1, so that the processes are also jointly wide-sense stationary. The cross-power spectrum of the processes, again shown in Appendix A.1, takes a form of

$$f_{xy}(\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_2)\exp(i(k-l)\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi}\left(1-\exp(i\lambda)\right)^{-d_1}\left(1-\exp(-i\lambda)\right)^{-d_2}. \tag{3.10}$$

To show whether the processes are long-range cross-correlated according to Definition 2.4, we need to inspect an asymptotic behavior of the cross-correlation function $\rho_{xy}(n)$. Using the inverse Fourier transform of the cross-power spectrum, we can write the nth cross-correlation as

$$\rho_{xy}(n) = \frac{\sigma_{\varepsilon\nu}}{2\pi\sigma_x\sigma_y} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_2) \int_{-\pi}^{\pi} \exp(i(n+k-l)\lambda)d\lambda. \tag{3.11}$$

Now, using the definition of Dirac delta function (Dirac 1958), we can rewrite the cross-correlation in Equation 3.11 as

$$\rho_{xy}(n) = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_2)\delta(n+k-l) = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_k(d_1)a_{n+k}(d_2) \tag{3.12}$$

For the second equality in Equation 3.12, we use the property of Dirac delta function

$$\int_0^{\infty} \delta(t-a)G(t)dt = G(a) \tag{3.13}$$

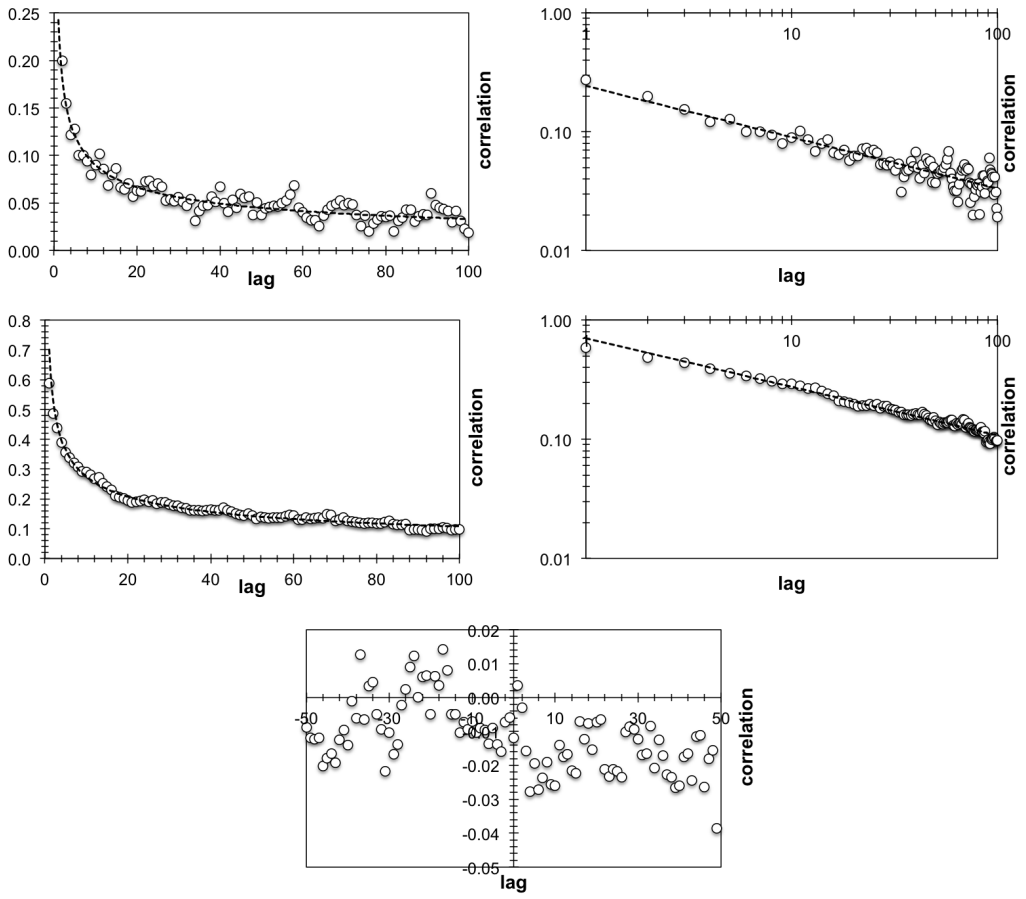


Figure 3.1: **Two ARFIMA processes with uncorrelated innovations.** Correlations structure for two ARFIMA processes with $d_1 = 0.2$ and $d_2 = 0.4$ with uncorrelated innovations. Power-law scaling of the auto-correlation functions is evident for the separate processes (the first two rows) while the cross-correlation function does not show any pattern (the last row).

for any continuous function $G(t)$ and its evenness so that

$$\sum_{l=0}^{\infty} a_k(d_1)a_l(d_2)\delta(n+k-l) = \sum_{l=0}^{\infty} a_k(d_1)a_l(d_2)\delta(l-n-k) = \sum_{l=0}^{\infty} a_k(d_1)a_{n+k}(d_2) \tag{3.14}$$

where $t = l$ and $a = n + k$. We can now rewrite $a_j(d)$ with a use of Beta function so that

$$a_j(d) = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} = \frac{1}{kB(k,d)}. \tag{3.15}$$

Using Stirling's approximation of Beta function $B(\bullet,\bullet)$ for fixed d and $j \to +\infty$, we get

$$a_j(d) \approx \frac{1}{j}\frac{1}{\Gamma(d)j^{-d}} = \frac{j^{d-1}}{\Gamma(d)}. \tag{3.16}$$

Since we are interested in the asymptotic behavior, we can use Equation 3.16 and follow with

$$\rho_{xy}(n) \approx \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y\Gamma(d_1)\Gamma(d_2)} \sum_{k=0}^{\infty} k^{d_1-1}(n+k)^{d_2-1} \tag{3.17}$$

given that $d_1, d_2, k, n+k > 0$. Approximating the infinite sum with a definite integral, we can write

$$\begin{aligned}\rho_{xy}(n) &\approx \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y\Gamma(d_1)\Gamma(d_2)} \int_0^{\infty} k^{d_1-1}(n+k)^{d_2-1}dk = \\ &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y\Gamma(d_1)\Gamma(d_2)} n^{d_1+d_2-1}\frac{\Gamma(d_1)\Gamma(1-d_1-d_2)}{\Gamma(1-d_2)} = \\ &= \frac{\sigma_{\varepsilon\nu}\Gamma(1-d_1-d_2)}{\sigma_x\sigma_y\Gamma(1-d_2)\Gamma(d_2)} n^{d_1+d_2-1} \propto n^{d_1+d_2-1} = n^{-(1-d_1-d_2)}\end{aligned} \tag{3.18}$$

given that $d_1 + d_2 < 1$ and $n > 0$. Therefore, given that $\sigma_{\varepsilon\nu} \neq 0$, the power-law cross-correlations emerge regardless of the level of correlation between innovations $\{\varepsilon_t\}$ and $\{\nu_t\}$. Using the relationship between fractional differencing parameter and Hurst exponent $d = H - 0.5$, and Definition 2.4, Equation 2.29 and 3.18, we have

$$H_{xy} = 1 - \frac{\gamma_{xy}}{2} = 1 - \frac{1-d_1-d_2}{2} = 1 - \frac{-(H_x+H_y)+2}{2} = \frac{H_x+H_y}{2}. \tag{3.19}$$

The bivariate Hurst exponent H_{xy} is thus an average of separate Hurst exponents H_x and H_y regardless the correlation between innovations. This also covers the case showed by Podobnik *et al.* (2009a) for two ARFIMA processes

with the identical innovations.

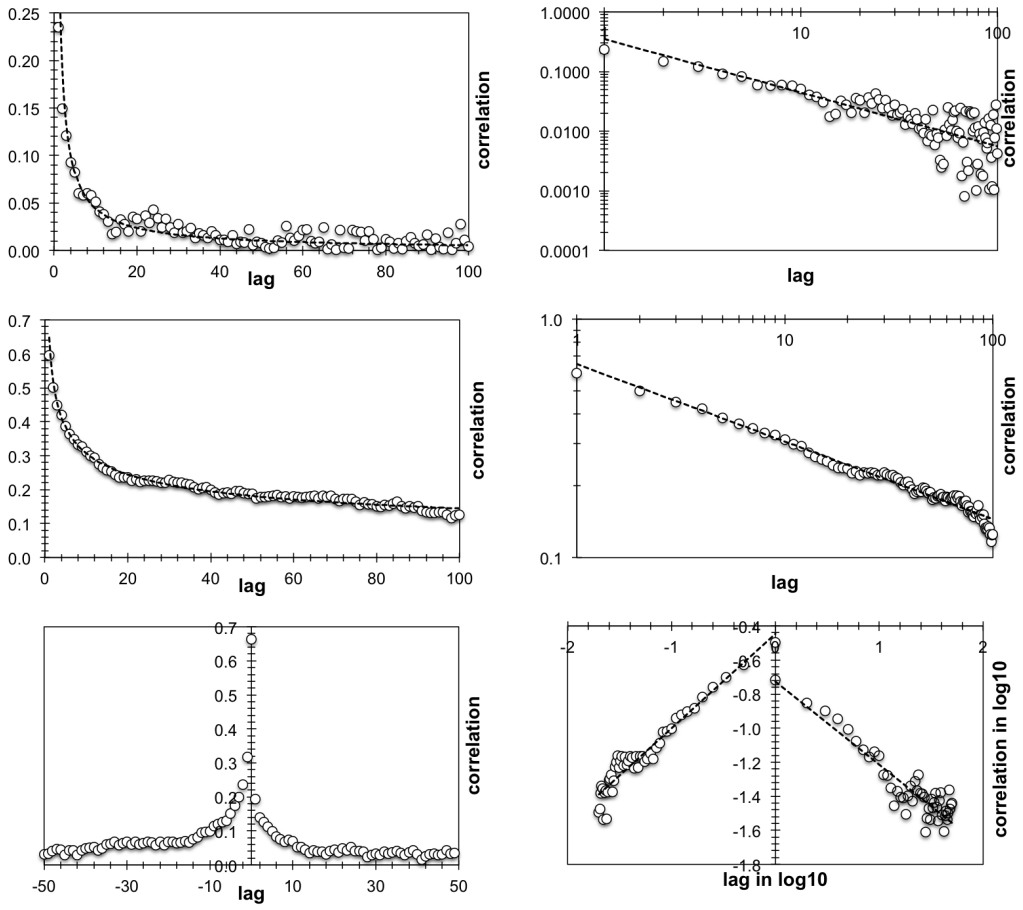


Figure 3.2: **Two ARFIMA processes with correlated innovations.** Correlations structure for two ARFIMA processes with $d_1 = 0.2$ and $d_2 = 0.4$ with correlated innovations, $\rho_{\varepsilon\nu} = 0.75$. Power-law scaling of the auto-correlation functions is evident for the separate processes (the first two rows) as well as for the cross-correlation function where we observe the hyperbolic decay for both positive and negative lags (the last row).

In Figure 3.1 and 3.2, we illustrate the behavior of the above analyzed cases. In Figure 3.1, we present the auto-correlation and cross-correlation functions for two ARFIMA processes with $d_1 = 0.2$, $d_2 = 0.4$, $\sigma_\varepsilon^2 = \sigma_\nu^2 = 1$ and $\sigma_{\varepsilon\nu} = 0$. For processes $\{x_t\}$ and $\{y_t\}$, we observe evident power-law scaling of the auto-correlation functions as a sign of long-term memory. The cross-correlation function contains values very close to zero for lags between -50 and 50 lags. Figure 3.2 illustrates the results for two ARFIMA processes with $d_1 = 0.2$, $d_2 = 0.4$, $\sigma_\varepsilon^2 = \sigma_\nu^2 = 1$ and $\sigma_{\varepsilon\nu} = 0.75$. Again, the auto-correlation functions of $\{x_t\}$ and $\{y_t\}$ show a hyperbolic decay representative for persistent processes. The cross-correlation functions can be nicely described as two power laws both

for positive and negative lags symbolic for the long-range cross-correlated processes.

In the same manner, it can be easily shown that the long-range cross-correlations arise for two long-range correlated processes which have innovations correlated at at least one finite lag or lead. For purposes of this thesis, it is sufficient to work with the simplest case of correlations between innovations as presented in this section.

3.2 ARFIMA(0,d,0) and AR(1) processes with correlated innovations

In the univariate case, distinguishing between short- and long-term memory is evident from the properties of the auto-correlation function. To see how these two types of memories interact in the bivariate setting, we investigate the case when one of the processes is long-range dependent, the other is short-range dependent and their innovations are pairwise correlated. Again for simplicity, we consider the case when the innovations are only correlated but not cross-correlated or auto-correlated. Let's have ARFIMA process $\{x_t\}$ and AR(1) process $\{y_t\}$ defined as

$$x_t = \sum_{n=0}^{\infty} a_n(d_1)\varepsilon_{t-n} \tag{3.20}$$

$$y_t = \theta y_{t-1} + \nu_t \tag{3.21}$$

with $|\theta| < 1$. The moments are specified as for the previous case (specified in Equation 3.4) – innovations have finite variance and the correlation structure is the simplest one with just pairwise correlated innovations.

As a starting point, we rewrite the AR(1) processes in the MA(∞) representation as

$$y_t = \theta y_{t-1} + \nu_t = \nu_t + \theta(\nu_{t-1} + \theta y_{t-2}) = \ldots = \sum_{n=0}^{\infty} \theta^n \nu_{t-n}. \tag{3.22}$$

Process $\{x_t\}$ is the same ARFIMA(0,d_1,0) as in the previous section and it has a zero mean and a finite variance and is thus stationary. Process $\{y_t\}$ described

in Equation 3.22 is evidently a zero mean process. Its variance can be found as

$$\mathrm{var}(y_t) = \langle y_t^2 \rangle = \langle (\nu_t + \theta\nu_{t-1} + \theta^2\nu_{t-2}\ldots)(\nu_t + \theta\nu_{t-1} + \theta^2\nu_{t-2}\ldots) \rangle = \sum_{j=0}^{\infty} \theta^{2j} \langle \nu_t^2 \rangle = \sigma_\nu^2 \sum_{j=0}^{\infty} \theta^{2j} = \frac{\sigma_\nu^2}{1-\theta^2} \equiv \sigma_y^2 < +\infty. \quad (3.23)$$

Auto-covariance and auto-correlation functions of $\{y_t\}$ are independent of t (Wei 2006) so that the process is stationary as well.

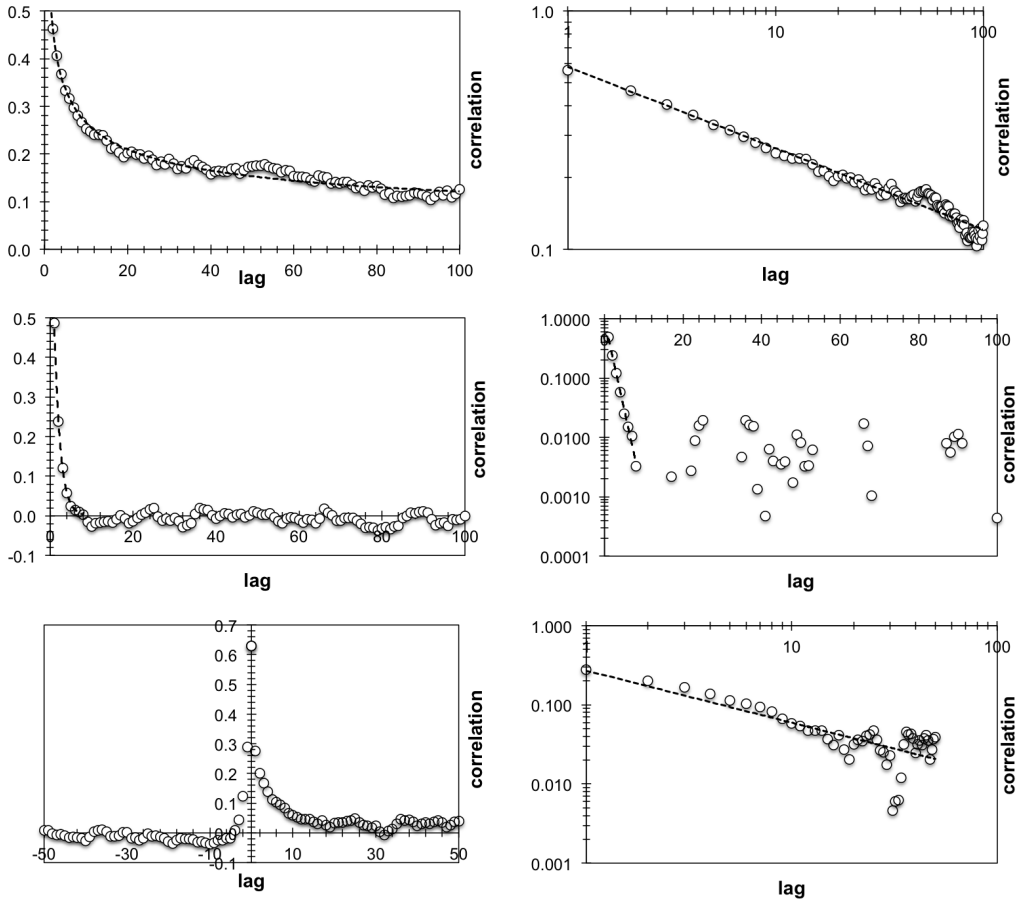


Figure 3.3: **ARFIMA and AR processes with correlated innovations.** Correlations structure for ARFIMA process with $d_1 = 0.4$ and AR(1) process with $\theta = 0.5$ with correlated innovations, $\rho_{\varepsilon\nu} = 0.75$. Power-law scaling of the auto-correlation functions is evident for the ARFIMA process (the first row), while the AR(1) process follows an exponential decay of auto-correlations (the second row). The cross-correlation function is strongly asymmetric and shows an evident power law decay for the positive lags (the third row).

The structure of cross-correlations is derived in Appendix A.2 and is shown to be independent of t so that the processes are also jointly stationary. Cross-

power spectrum, derived in Appendix A.2, has the following form

$$f_{xy}(\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_l(d_1)\theta^k \exp(i(k-l)\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \left(1 - \exp(-i\lambda)\right)^{-d_1} \left(1 - \theta \exp(i\lambda)\right)^{-1}. \quad (3.24)$$

Using the inverse Fourier transform, we get

$$\rho_{xy}(n) = \frac{\sigma_{\varepsilon\nu}}{2\pi\sigma_x\sigma_y} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_l(d_1)\theta^k \int_{-\pi}^{\pi} \exp(i(n+k-l)\lambda) d\lambda. \quad (3.25)$$

Again, we use the definition of Dirac's delta function and its properties to get

$$\rho_{xy}(n) = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_l(d_1)\theta^k \delta(n+k-l) = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_{n+k}(d_1)\theta^k. \quad (3.26)$$

Using the Stirling's approximation and approximating the infinite sum by the definite integral, we get

$$\rho_{xy}(n) \propto \int_0^{\infty} (n+k)^{d_1-1}\theta^k dk = \theta^{-n}\Gamma(d_1, -n\log\theta)(-\log\theta)^{-d_1} \quad (3.27)$$

where $\Gamma(\bullet, \bullet)$ is the incomplete upper Gamma function (Wall 1948). Using the approximation of the incomplete upper Gamma function in Blahak (2010), we can write

$$\rho_{xy}(n) \propto \theta^{-n}(-\log\theta)^{-d_1}(-n\log\theta)^{d_1-1} \exp(n\log\theta) = \theta^{-n}\theta^n(-\log\theta)^{-d_1}(-\log\theta)^{d_1+1} n^{d_1-1} \propto n^{d_1-1} \quad (3.28)$$

Therefore, we have

$$H_{xy} = 1 - \frac{\gamma_{xy}}{2} = 1 - \frac{1-d_1}{2} = 1 - \frac{-H_x + 1.5}{2} = \frac{H_x + 0.5}{2} \quad (3.29)$$

which is perfectly in hand with Equation 3.19 for $H_y = 0.5$, i.e. the process $\{y_t\}$ is not long-range dependent with $d_2 = 0$. Note that the asymptotic relationship is again independent of $\sigma_{\varepsilon\nu}$ as long as $\sigma_{\varepsilon\nu} \neq 0$.

In Figure 3.3, we present the simulations results for $T = 10000$, $d_1 = 0.4$, $\theta = 0.5$, $\sigma_\varepsilon^2 = \sigma_\nu^2 = 1$ and $\sigma_{\varepsilon\nu} = 0.75$. The auto-correlation functions show the obvious difference between the short- and long-range dependence – the long-

range dependent $\{x_t\}$ follows a power-lay decay in auto-correlations while the auto-correlations of the short-range dependent $\{y_t\}$ decay exponentially for the first several lags until reaching the noise level. This is reflected in the cross-correlation function which shows a power-law decay for the positive lags but vanishes rapidly for the negative lags.

3.3 Mixed-correlated ARFIMA(0,d,0) processes

In the previous two sections, we have shown two types of processes leading to the bivariate Hurst exponent which is equal to the average of the separate Hurst exponents. In this section, we present a new kind of bivariate process which allows for manipulation of H_{xy}, H_x and H_y as long as the bivariate Hurst exponent is not greater than the average of the separate Hurst exponents. Let's have two processes

$$\begin{aligned} x_t &= \alpha \sum_{n=0}^{\infty} a_n(d_1)\varepsilon_{1,t-n} + \beta \sum_{n=0}^{\infty} a_n(d_2)\varepsilon_{2,t-n} \\ y_t &= \gamma \sum_{n=0}^{\infty} a_n(d_3)\varepsilon_{3,t-n} + \delta \sum_{n=0}^{\infty} a_n(d_4)\varepsilon_{4,t-n}. \end{aligned} \tag{3.30}$$

Innovations are characterized by

$$\begin{aligned} \langle \varepsilon_{i,t} \rangle &= 0 \text{ for } i = 1,2,3,4 \\ \langle \varepsilon_{i,t}^2 \rangle &= \sigma_{\varepsilon_i}^2 \text{ for } i = 1,2,3,4 \\ \langle \varepsilon_{i,t}\varepsilon_{j,t-n} \rangle &= 0 \text{ for } n \neq 0 \text{ and } i,j = 1,2,3,4 \\ \langle \varepsilon_{i,t}\varepsilon_{j,t} \rangle &= \sigma_{ij} \text{ for } i,j = 1,2,3,4 \text{ and } i \neq j. \end{aligned} \tag{3.31}$$

In words, we have two processes and each one is a linear combination of two long-range correlated processes with possibly correlated innovations. Note that the separate long-term memory parameters d_1, d_2, d_3, d_4 can vary or be the same. We call the set of processes $\{x_t\}$ and $\{y_t\}$ as the mixed-correlated ARFIMA processes (MC-ARFIMA). As MC-ARFIMA is a new kind of process not discussed in the literature, even though these can be seen as a generalization of the anti-cointegration model of Sela & Hurvich (2012), we need to at least shortly discuss its stationarity and joint-stationarity. For the wide-sense stationarity, it suffices to state that both $\{x_t\}$ and $\{y_t\}$ are linear combinations of two ARFIMA(0,d,0) processes with correlated innovations which are wide-sense

stationary as shown in Section 3.1 so that MC-ARFIMA processes are stationary as long as $0 \leq d_1, d_2, d_3, d_4 < 0.5$. Evidently, we have $\langle x_t \rangle = \langle y_t \rangle = 0$ and both processes have finite variance, i.e. $\langle x_t^2 \rangle \equiv \sigma_x^2 < +\infty$ and $\langle y_t^2 \rangle \equiv \sigma_y^2 < +\infty$. As ARFIMA(0,$d$,0) processes are long-range correlated, their linear combination is also long-range correlated. The higher d will dominate in the linear combination so that process $\{x_t\}$ is integrated of order $\max(d_1, d_2)$ and $\{y_t\}$ of order $\max(d_3, d_4)$. The processes are thus wide-sense stationary.

To show that $\{x_t\}$ and $\{y_t\}$ are also jointly wide-sense stationary, we need to show that $\rho_{xy}(k)$ does not depend on t. As shown in Appendix A.3, the cross-correlation function is dependent only on the parameters d_1, d_2, d_3, d_4, α, β, γ, δ and σ_{ij} (with $i, j = 1, 2, 3, 4$) and the processes $\{x_t\}$ and $\{y_t\}$ are thus also jointly wide-sense stationary.

The cross-power spectrum is derived in Appendix A.3 as well, and it can be written as

$$\begin{aligned} f_{xy}(\lambda) = \frac{\alpha\gamma\sigma_{13}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_3)\exp(i(k-l)\lambda) + \\ \frac{\alpha\delta\sigma_{14}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_4)\exp(i(k-l)\lambda) + \\ \frac{\beta\gamma\sigma_{23}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_2)a_l(d_3)\exp(i(k-l)\lambda) + \\ \frac{\beta\delta\sigma_{24}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_2)a_l(d_4)\exp(i(k-l)\lambda) = \\ \frac{1}{2\pi}\Big[\alpha\gamma\sigma_{13}\left(1-\exp(i\lambda)\right)^{-d_1}\left(1-\exp(-i\lambda)\right)^{-d_3} + \\ \alpha\delta\sigma_{14}\left(1-\exp(i\lambda)\right)^{-d_1}\left(1-\exp(-i\lambda)\right)^{-d_4} + \\ \beta\gamma\sigma_{23}\left(1-\exp(i\lambda)\right)^{-d_2}\left(1-\exp(-i\lambda)\right)^{-d_3} + \\ \beta\delta\sigma_{24}\left(1-\exp(i\lambda)\right)^{-d_2}\left(1-\exp(-i\lambda)\right)^{-d_4}\Big]. \end{aligned} \tag{3.32}$$

Following the same steps as in the previous cases, we use the inverse Fourier

transform and the Dirac delta function to get

$$\begin{aligned}\rho_{xy}(n) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_k(d_1)a_l(d_3)\delta(n+k-l)+ \\ \frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_k(d_1)a_l(d_4)\delta(n+k-l)+ \\ \frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_k(d_2)a_l(d_3)\delta(n+k-l)+ \\ \frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_k(d_2)a_l(d_4)\delta(n+k-l) = \\ \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\underbrace{\sum_{k=0}^{\infty}a_k(d_1)a_{n+k}(d_3)}_{\approx\int_0^\infty k^{d_1-1}(n+k)^{d_3-1}dk\propto n^{d_1+d_3-1}} + \frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\underbrace{\sum_{k=0}^{\infty}a_k(d_1)a_{n+k}(d_4)}_{\approx\int_0^\infty k^{d_1-1}(n+k)^{d_4-1}dk\propto n^{d_1+d_4-1}} + \\ \frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\underbrace{\sum_{k=0}^{\infty}a_k(d_2)a_{n+k}(d_3)}_{\approx\int_0^\infty k^{d_2-1}(n+k)^{d_3-1}dk\propto n^{d_2+d_3-1}} + \frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\underbrace{\sum_{k=0}^{\infty}a_k(d_2)a_{n+k}(d_4)}_{\approx\int_0^\infty k^{d_2-1}(n+k)^{d_4-1}dk\propto n^{d_2+d_4-1}}. \quad (3.33)\end{aligned}$$

The results are obtained by using the Stirling's approximation and by approximating the infinite sum by the definite integrals. As we are interested in the asymptotic case $n \to +\infty$, the scaling of $\rho_{xy}(n)$ will be dominated by the highest exponent. This leads us to several interesting settings arriving at various behaviors of the bivariate Hurst exponent.

First, let's have $\alpha, \beta, \gamma, \delta \neq 0$ and $\sigma_{ij} \neq 0$ for all $i, j = 1, 2, 3, 4$. Labeling $H_i = d_i + 0.5$ for $i = 1, 2, 3, 4$, we have

$$\begin{aligned}H_x &= \max(H_1, H_2) \\ H_y &= \max(H_3, H_4) \quad (3.34)\end{aligned}$$

$$\begin{aligned}H_{xy} = \frac{\max(H_1+H_2, H_1+H_4, H_2+H_3, H_2+H_4)}{2} = \\ \frac{\max(H_1,H_2)+\max(H_3,H_4)}{2} = \frac{H_x+H_y}{2}, \quad (3.35)\end{aligned}$$

which is the same result that we have obtained several times before.

Second, let's again have $\alpha, \beta, \gamma, \delta \neq 0$ and without loss on generality, let's have $\max(H_1, H_2) = H_1$ and $\max(H_3, H_4) = H_4$ so that $\{x_t\}$ is integrated of order d_1 and $\{y_t\}$ of order d_4. Moreover, assume that $\sigma_{23} = \sigma_{32} \neq 0$ and all the

other covariances are equal to zero. From Equation 3.33, this implies

$$H_{xy} = \frac{H_2 + H_3}{2} < \frac{H_x + H_y}{2} = \frac{H_1 + H_4}{2} = \frac{\max(H_1, H_2) + \max(H_3, H_4)}{2}. \tag{3.36}$$

This is the situation when the bivariate Hurst exponent H_{xy} is not equal to the average of the univariate Hurst exponents H_x and H_y while still showing long-range cross-correlations, i.e. without $H_{xy} = 0.5$, which has been shown for example for two ARFIMA(0,d,0) processes with uncorrelated innovations. Processes $\{x_t\}$ and $\{y_t\}$ are thus long-range cross-correlated but possess the power law coherency in a similar manner as the anti-cointegration model of Sela & Hurvich (2012).

Third, let's set $\beta = 0$, $0 < d_4 < d_1 = d_3$ and $\sigma_{13} = \sigma_{31} = \sigma_{\varepsilon_1}\sigma_{\varepsilon_3}$, i.e. the innovations processes $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{3t}\}$ are identical. Therefore, the setting in Equation 3.30 reduces to the fractional cointegration case (Baillie & Bollerslev 1994; Gil-Alana & Hualde 2009) with $d = d_1 = d_3$ and $d_U = d_4$. It thus again holds that $H_x = H_y = H_{xy}$ and from the LRCC point of view, the fractional cointegration relationship is not different from the first case discussed in this subsection.

Fourth, let's have the same setting as in the previous case but additionally, let's have $H_2 = H_3 = 0.5$. Reformulating Equation 3.30 yields

$$\begin{aligned} x_t &= \alpha \sum_{n=0}^{\infty} a_n(d_1)\varepsilon_{1,t-n} + \beta\varepsilon_{2,t} \\ y_t &= \gamma\varepsilon_{3,t} + \delta \sum_{n=0}^{\infty} a_n(d_4)\varepsilon_{4,t-n}. \end{aligned} \tag{3.37}$$

We thus again have $\{x_t\}$ with the long-term memory parameter d_1 and $\{y_t\}$ with d_4 but $H_{xy} = 0.5$ and according to Appendix A.3, we have $\rho_{xy}(0) = \frac{\sigma_{23}}{\sigma_x\sigma_y}$ and $\rho_{xy}(k) = 0$ for $k \neq 0$. Therefore, we have two long-range dependent processes which are correlated but not cross-correlated. This is nicely documented in Figure 3.4 where we show statistical properties of MC-ARFIMA processes with $\alpha = \beta = \gamma = \delta = 1$, $d_1 = d_4 = 0.4$, $\sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2} = \sigma^2_{\varepsilon_3} = \sigma^2_{\varepsilon_4} = 1$, $\sigma_{23} = \sigma_{32} = 0.5$ and other covariances are equal to zero, with $T = 10000$. It is shown that both $\{x_t\}$ and $\{y_t\}$ are evidently long-range correlated as their auto-correlation functions are well described with a power-law decay. On contrary, the cross-correlation function has a non-zero value at lag zero and insignificant values at other lags.

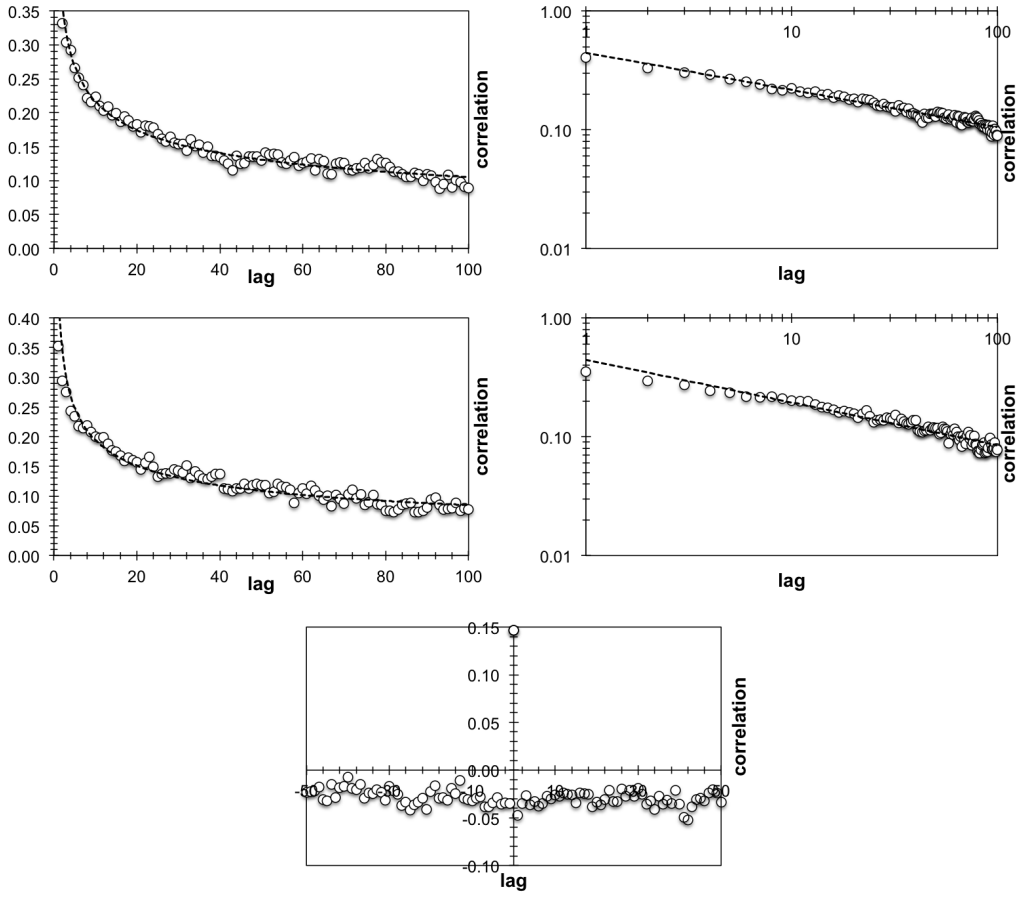


Figure 3.4: **Mixed-correlated ARFIMA processes.** MC-ARFIMA specified by $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0$ with correlated innovations, $\rho_{\varepsilon_2\varepsilon 3} = 0.5$. The separate processes follow the power-law decay of auto-correlation function (the first two rows) while the cross-correlation function shows just pairwise correlated processes.

And fifth, let's have the same assumptions about correlations and parameters $\alpha, \beta, \gamma, \delta$ as in the previous two cases but let's adjust the model specification to

$$x_t = \alpha \sum_{n=0}^{\infty} a_n(d_1)\varepsilon_{1,t-n} + \beta \sum_{n=0}^{\infty} \theta_2^n \varepsilon_{2,t-n}$$
$$y_t = \gamma \sum_{n=0}^{\infty} \theta_3^n \varepsilon_{3,t-n} + \delta \sum_{n=0}^{\infty} a_n(d_4)\varepsilon_{4,t-n}. \tag{3.38}$$

Processes $\{x_t\}$ and $\{y_t\}$ are thus linear combinations of ARFIMA(0,d,0) and AR(1) processes. In this case, we again have $\{x_t\}$ with memory d_1 and $\{y_t\}$ with memory d_4. And as the only non-zero correlation between innovations is $\sigma_{23} = \sigma_{32}$, we have, based on Equation 3.33 and Section 3.2, $H_{xy} = 0.5$. We thus have two long-range correlated processes $\{x_t\}$ and $\{y_t\}$, which are only short-range cross-correlated. Note that by setting $\alpha = \delta = 0$, we arrive at the correlated AR(1) processes so that we arrive at short-range correlated $\{x_t\}$ and $\{y_t\}$ which are also short-range cross-correlated.

The MC-ARFIMA framework thus provides quite a wide range of possible model specifications yielding long-range cross-correlated processes which can be either long-range cross-correlated (with or without power law coherency), short-range cross-correlated, pairwise correlated or uncorrelated. Generally, the framework allows for even more possible specifications.

3.4 Brief overview

We have shown that various specifications of processes can yield hyperbolically decaying cross-correlation function. Even quite simply defined ARFIMA processes with correlated innovations possess the power-law cross-correlations. These cross-correlations are thus just a result of the fact that the processes are separately long-range correlated and their innovations are correlated so that there is no causal relationship behind the cross-correlations. Such interpretation of slowly decaying cross-correlations must be thus taken with caution. Apparently, the power-law scaling of cross-correlations arises between the series if at least one of them is long-range correlated and the innovations of the processes are correlated at at least one lag or lead. Even though such a relationship is asymptotic, it shows that the situation when the series are long-range

cross-correlated with the bivariate Hurst exponent being equal to the average of the separate Hurst exponents can arise almost out of nowhere.

The second presented bivariate process – the combination of ARFIMA(0,d,0) and AR(1) processes with correlated innovations – nicely illustrates that a seemingly impulse–response relationship between two processes can easily arise from this specification. Resulting cross-correlation function is strongly asymmetric and obeys an evident power-law decay in one part of the function while in the other, the cross-correlations fall rapidly to the noise level as expected for the exponentially vanishing short-term correlated process. Nonetheless, the specification leads to the bivariate Hurst exponent equal to the average of the separate Hurst exponents.

The previous two cases show that the cross-persistent series can arise quite easily and very importantly, these could easily occur in macroeconomics and finance. Specifically for finance or financial economics, there are several processes with long-term memory which are well-documented in the literature – volatility, volume and signs of changes to name the most evident ones (Bollerslev & Jubinski 1999; Thomakos & Wang 2003; Poon & Granger 2003; 2005; Chen *et al.* 2006; Forsberg & Ghysels 2007; Fleming & Kirby 2011). However, when realizing that practically all processes in a specific class have a common information set, their innovations will be at least somehow (even if very weakly) correlated. This means that for example volatility processes of even remotely related stocks will be long-range cross-correlated. Evidently, this relationship would not be causal even though the cross-correlation structure might appear that way. This possibility is discussed in detail in Chapter 6 where we deal with relationship between stock index returns and realized volatility.

Apart from the cases when the bivariate Hurst exponent is equal to the average of the separate Hurst exponents, we introduced the mixed-correlated ARFIMA processes which are quite flexible for mixing of various types of memories in and between the processes. This kind of processes is important for the later Monte Carlo simulations study showing the statistical properties of various estimators of the bivariate Hurst exponent. Importantly, the MC-ARFIMA specification allows for the power law coherency.

Chapter 4

Tests for long-range cross-correlated processes

Long-range cross-correlated processes are characteristic by the power-law decay of the cross-correlation function, the power-law scaling of the covariances of the partial sums and the divergent at origin magnitude of the cross-power spectrum as we have shown in the previous sections. However, it has been also hinted that some of these properties might not be easy to check empirically as the properties are asymptotic. Even for the univariate case of the long memory, it has been frequently shown that the long-range correlations might be spuriously found for the processes which are only short-range correlated (Lo 1991; Kristoufek 2012) or due to distributional properties of the series (Barunik & Kristoufek 2010). To control for these effects, several tests for long-range correlations have been proposed in the literature. Lo (1991) constructs a test based on the modified rescaled range analysis which is robust to the short-range dependence bias. Even though its statistical properties have been discussed and disputed (Teverovsky *et al.* 1999; Kristoufek 2010a; 2012), it is still frequently used in the literature. As an alternative to the previous test, Giraitis *et al.* (2003) propose the rescaled variance test which they show to have better statistical properties than the previous test.

For the long-range cross-correlated processes, there has been, up to our knowledge, only one test proposed. The test based on the detrended cross-correlation analysis has been proposed by Zebende (2011) and further discussed by Podobnik *et al.* (2011). We present this test as the first one in this chapter. We then develop three new tests for a presence of the long-range cross-correlations between two series. For each test, we discuss its statistical

properties, namely its size and power. Since the alternative hypothesis of the long-range cross-correlation can arise from various models and specifications (as shown in the previous Chapter 3), we utilize Monte-Carlo-based sizes and powers of the tests as initial measures of performance of the tests.

Note that testing of processes being long-range cross-correlated should be applied before an estimation of the bivariate Hurst exponent or power law coherency because, as it is shown in the next chapter, some of the estimators have quite high variance or are even biased for finite samples which might lead to spurious conclusions even if the processes are not long-range cross-correlated at all.

4.1 Detrended cross-correlation coefficient

Zebende (2011) proposes to use a combination of DCCA and DFA, which are introduced in Subsection 5.1.1, to construct a correlation coefficient for detrended series, which might be also asymptotically non-stationary (with $H \geq 1$), as

$$\rho_{DCCA}(s) = \frac{F^2_{DCCA}(s)}{F_{DFA,x}(s)F_{DFA,y}(s)} \tag{4.1}$$

where $F^2_{DCCA}(s)$ is a detrended covariance between series $\{X_t\}$ and $\{Y_t\}$ for a window size s, and $F^2_{DFA,x}$ and $F^2_{DFA,y}$ are detrended variances of series $\{X_t\}$ and $\{Y_t\}$, respectively, for a window size s. Detrended variances can be seen as detrended covariances between two identical series. The basic idea of the test is based on scaling of detrended covariance. If the processes are long-range cross-correlated, the test statistic converges to a constant for high scales s. The constant itself depends on a specification of the process. The statistics $\rho_{DCCA}(s)$ goes to 1 for fractionally cointegrated processes, to 0 for anti-cointegrated processes, or processes with power law coherency properties in general, or the statistic goes to a constant between -1 and 1 for other cases of long-range cross-correlations. For short-range cross-correlated processes, the statistic should vanish quickly to zero because of the dominance of the denominator. However, this is only true for the case when the separate processes are long-range correlated but not long-range cross-correlated and it is thus not evident how the test fares in the case of the short-range correlated and cross-correlated processes. We shall see in the following that this is the most severe weakness of the test.

Podobnik *et al.* (2011) further explore statistical properties of the $\rho_{DCCA}(s)$ statistic. They analytically show that for the non-overlapping boxes, the statistic converges to the Gaussian distribution with zero mean and a variance of T^{-1} regardless of the box size s. For the overlapping windows, the statistical properties are more complicated and are not shown analytically but only numerically. Therefore, the two-sided critical values are simply equal to $\pm\Phi(\frac{\alpha}{2})/\sqrt{T}$, where $\Phi(\bullet)$ is a cumulative distribution function of the standard Gaussian distribution, for a given significance level α. We can thus easily use the statistic for testing whether the underlying series are long-range cross-correlated by using the test statistic with non-overlapping windows.

To evaluate the ability of $\rho_{DCCA}(s)$ statistic to uncover long-range cross-correlated series, we calculate its size and power for processes under the null hypothesis (short-range cross-correlated processes) and under the alternative hypothesis (long-range cross-correlated processes) in the Monte Carlo simulations setting. As the confidence intervals for the null hypothesis are simply $\pm\Phi(\frac{\alpha}{2})/\sqrt{T}$, we can set a specific level α and check whether this proportion is matched in the simulations. The size of the test is the proportion of the times when the null hypothesis is rejected while it is actually true, i.e. in the ideal case, the size of the test is equal to α. The power of the test, reversely, is the proportion of the times when the null hypothesis is rejected while the alternative hypothesis is actually true. Theoretically, the lower the size of the test the better and the higher the power of the test the better.

We are mainly interested in how the statistic performs under varying strength of correlation and cross-correlations, varying s, varying T and varying type of cross-correlations (no, short or long). In order to do so, we construct processes with differently correlated innovations ($\rho = 0, 0.5, 0.9$) and we estimate $\rho_{DCCA}(s)$ for $s = \frac{T}{100}, \frac{T}{50}, \frac{T}{10}, \frac{T}{5}$ with time series length $T = 500, 1000, 5000$. To control for a simple correlation between the processes, we construct simple correlated noise series. For short-range cross-correlated processes, we simulate AR(1) processes with $\theta_1 = \theta_2 = 0.1, 0.5, 0.8$, to control for weak, medium and strong short-term memory, and with correlated innovations. And for long-range cross-correlated processes, we analyze ARFIMA(0,d,0) processes with correlated innovations and with $d_x = d_y = 0.1, 0.4$ to control for weak and strong long-term memory.

Table B.1 presents the results for simply correlated processes with no cross-correlations. We observe that the size of the test never matches the theoretical α. The test gets the closest to the theoretical values for the shortest time series

($T = 500$) and the shortest box size ($s = 5$) of the uncorrelated processes, where we observe the sizes of 0.047, 0.145 and 0.241 for the significance levels of 0.01, 0.05 and 0.10, respectively. For the uncorrelated processes, the size of the test increases with the time series length T and with the box size s. Even for the uncorrelated processes, the test practically breaks down for the series of $T = 5000$. The situation gets critical for the series which are correlated. For both $\rho = 0.5$ and $\rho = 0.9$, the size of the test approaches 1, i.e. the test practically always rejects the null hypothesis. Such results are well against the intuitive size of the test, which we expected to decrease with increasing T and s (based on an assumption that the long-range cross-correlations would be observed at higher scales s). Presented sizes for correlated processes go against the analytical results presented by Podobnik *et al.* (2011).

Practically the same behavior of the test sizes are observed for short-range cross-correlated processes (Table B.2-B.4). The sizes are well above the theoretical ones and in fact, they give at least some information only for the processes with uncorrelated innovations, i.e. two short-range dependent processes which are uncorrelated. Again, the size of the test increases with T and s, which is undesirable. For cross-correlated processes with $\rho = 0.5$ and $\rho = 0.9$, the test breaks down regardless the strength of the short-range correlations. The test thus confuses short-range cross-correlations with long-range cross-correlations almost every time.

Based on the previous results, we might expect that the test detects the long-range cross-correlations practically always, which is indeed the case. In Table B.5-B.6, we can see that apart from uncorrelated ARFIMA processes, which show similar sizes to the other uncorrected cases discussed previously, the test rejects the null hypothesis of short-range cross-correlations practically every time the processes are in fact long-range cross-correlated. Such results would be highly desirable if considered solely. However, when we summarize the size and power of the test, we conclude that the test is practically useless for detecting the long-range cross-correlations as it finds them almost always regardless the type of cross-correlations present in the processes.

4.2 Aggregate cross-correlations test

As we have shown in Proposition 2.1 and Proposition 2.2, the sum of cross-correlations of the long-range cross-correlated processes diverges while the sum converges for the short-range cross-correlated processes. Even though these are

asymptotic properties, it is reasonable to assume that the divergence and convergence of the sums would be pronounced for finite or even for relatively low, lags. To illustrate such a statement, we compare behavior of cross-correlation functions for weak and strong short- and long-range cross-correlated processes. For the short-range cross-correlated processes, we show simulated series for two AR(1) processes with perfectly correlated innovations, and $\theta = 0.1$ for weak memory and $\theta = 0.9$ for strong memory. For the long-range cross-correlated processes, we simulate two ARFIMA(0,d,0) processes with perfectly correlated innovations, and $d = 0.1$ for weak and $d = 0.4$ for strong memory. 10,000 realizations are simulated for each specification. Figure 4.1 shows the cross-correlations in the semi-logarithmic representation for better comparison.

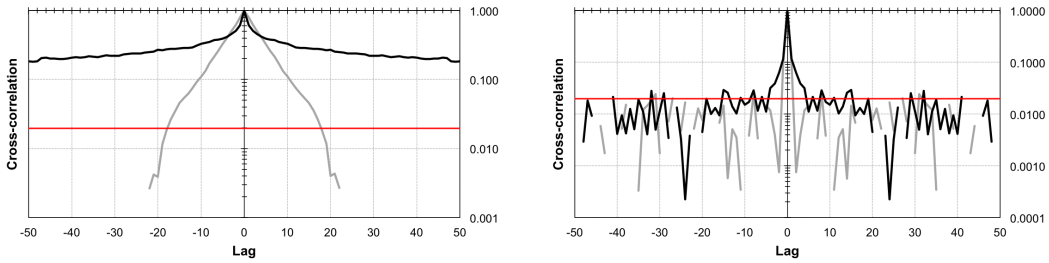


Figure 4.1: **Cross-correlation functions for short- and long-range cross-correlated processes.** On the left chart, a comparison of strong short- and long-term cross-correlations with AR(1) process with $\theta = 0.9$ (grey) and ARFIMA(0,0.4,0) (black), both with perfectly correlated innovations is shown. On the right chart, a comparison of weak short- and long-term cross-correlations with AR(1) process with $\theta = 0.1$ (grey) and ARFIMA(0,0.1,0) (black), both with perfectly correlated innovations is shown. Red line represents a 95% critical level for uncorrelated processes. Semi-logarithmic depiction is showed for better comparison (negative cross-correlations are in effect not shown).

For strong levels of memory, we observe that for the short-term dependence, the cross-correlations decay exponentially and vanish below the noise level after approximately 18 lags. On contrary, the cross-correlations of the long-term dependent processes decay slowly and do not get even close to the noise level for 50 lags which are presented in the chart. The difference between these two types of processes is thus evident even for rather short lags. Moreover, it should be noted that the strength of short-term memory is higher than the strength of long-term memory in a way that $\theta = 0.9$ is much closer to a unit-root case ($\theta = 1$) than the case of $d = 0.4$ (with respect to a unit root of $d = 1$). For weak levels of memory, the difference is obviously not so pronounced. Nonetheless, the AR(1) specification shows significant cross-

correlations only for the first lag and these then vanish below the noise level while the ARFIMA(0,0.1,0) specification brings significant cross-correlations up to the fourth lag and these fluctuate around the noise level even for higher lags. Based on the presented different behavior of the cross-correlation function for short- and long-range cross-correlated processes even for finite lags, we propose a test for distinguishing between the two types of memories.

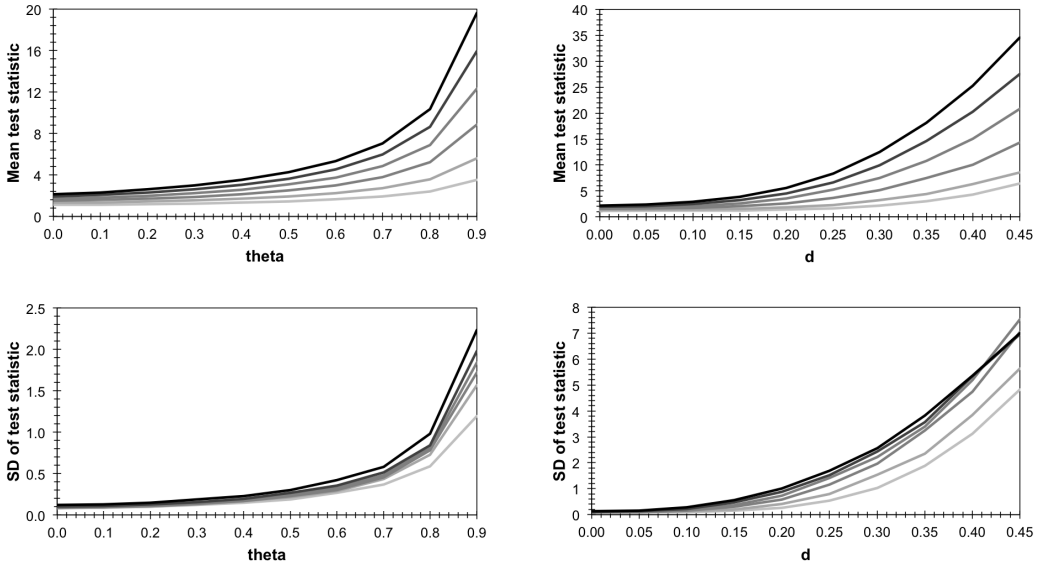


Figure 4.2: **Mean values and standard deviations of ACC test.** Test statistic ξ_{50} for differently correlated processes. Correlation between innovations varies between 0 and 1 with a step of 0.2 and the darker the line in the chart is, the higher the correlation is. On the left, correlated AR(1) processes with θ ranging between 0 and 0.9 with a step of 0.1 are shown. On the right, correlated ARFIMA(0,d,0) processes with d ranging between 0 and 0.45 with a step of 0.05 are shown. Means are based on 1,000 simulations with a time series length of 5,000.

Definition 4.1 (Aggregate cross-correlations test). Let processes $\{x_t\}$ and $\{y_t\}$, with $t = 1, \ldots, T$, be jointly wide-sense stationary processes. An aggregate cross-correlations statistic is defined as

$$\xi_k = \sum_{i=-k}^{k} |\rho_{xy}(i)|. \tag{4.2}$$

The null and alternative hypotheses are stated as "short-range cross-correlations" and "long-range cross-correlations", respectively. As we are dealing with series of a finite length, the lag parameter k is always finite as well. The testing statistic ξ_k is thus evidently dependent on k. Moreover, as we have shown

in Proposition 2.2, the sum of cross-correlations is also dependent on their strength (characterized by parameter δ in the proposition). Moreover, ξ_k is also dependent on the actual level of correlations and not only on their decay. The test statistic will thus not converge to a unique value even for the null hypothesis. This is nicely illustrated in Figure 4.2 which illustrates the behavior of a mean value of ξ_k statistic for $k = 50$ based on 1,000 simulations of short- and long-range cross-correlated processes with 5,000 observations. We observe that the testing statistic increases with θ and d as expected and also with the correlation between innovations of the cross-correlated processes, which is not surprising either. It is also evident that the ξ_{50} statistic approaches much higher values for the long-range cross-correlated processes than for the short-term dependent ones, which is well in hand with the asymptotic properties of the sum of cross-correlations. The variance of the test statistic is evidently dependent on both the strength of memory (θ and d) and correlations between innovations. Even though the levels of variance are higher for the long-term memory case, these do not deviate strongly from the short-term memory processes.

Nonetheless, the values of the testing statistic are evidently dependent on various parameters even under the null hypothesis. To control for such dependence, we use the moving-block bootstrap (MBB) method to construct confidence intervals and to obtain p-values. In the process, one obtains a bootstrapped series by separating the series into blocks of size ζ and shuffling the blocks, then the parameter of interest is estimated on the bootstrapped series which retains the short-range dependence and the distributional properties of the original series. After B repetitions of MBB and obtaining B estimates of the parameter of interest, we obtain empirical confidence intervals for a specific level α and an empirical p-value. For more detailed treatment of bootstrapping methods and the moving block bootstrap in particular, see Efron (1979), Efron *et al.* (1993), Davison & Hinkley (1997) and Srinivas & Srinivasan (2000). In effect, we have a one-sided test with the null hypothesis of short-range cross-correlated processes against the alternative hypothesis of cross-persistence.

In Tables B.7-B.12, we present the power and size of the test for the same processes as for the detrended cross-correlation coefficient in the previous section – correlated noise, correlated AR(1) processes and correlated ARFIMA processes. The statistical properties are analyzed for various numbers of lags taken into consideration for the test – $k = 5, 10, 25, 50$ – with $\zeta = 25$. For correlated noise, we can see that the size of the ACC test matches the significance levels for practically all cases, regardless the time series length, number of lags

taken into consideration or the strength of correlation (Table B.7). For short-range cross-correlated processes, the size of the test deteriorates a little. For weak correlations ($\theta = 0.1$), the situation is hardly distinguishable for most of the cases with $k \geq 10$. For medium correlations ($\theta = 0.5$), the results worsen. However, even for $\rho = 0.9$, the size of the test is close to the theoretical values if enough lags are taken into consideration. Interestingly, the statistical size worsens with the time series length. For very strong correlations ($\theta = 0.8$), the test does not perform well unless for a very short series. It would be needed to take more lags into consideration to make the test more robust. Generally for the short-range dependent processes, the size of the test is evidently dependent on the number of lags taken into consideration. Lags k need to be increased with an increasing memory parameter θ.

As for the power of the test, Table B.11 shows that for weakly long-range cross-correlated processes ($d = 0.1$), the test is able to capture the memory if the level of correlations and the time series length are high. For shorter series, the test is not able to capture the memory sufficiently. For strong long-range cross-correlations ($d = 0.4$), the test is well able to detect the long memory. For both levels of correlations and all time series lengths, the power of the test is high. However, when the innovations are not correlated, which implies that the processes are not long-range cross-correlated, the test quite often identifies the processes as long-range cross-correlated. Therefore, if the tested processes are not at all correlated, the test does not perform that well, which needs to be kept in mind. The power of the test decreases with number of lags taken into consideration, which is indeed expected as the more lags are used, the more probable it is that the test confuses the long-memory with short-memory. However, if the tested series are long (in our case $T = 5000$), the number of lags does not influence the ability to detect long-range cross-correlations.

4.3 Partial sums covariance divergence test

Similarly to the previous test, we base the partial sums covariance divergence test (PSCD) on one of the propositions introduced in Chapter 2. Proposition 2.4 states that the covariance of partial sums of the long-range cross-correlated processes divided by the time series length diverges as the time series length goes to infinity while the same ratio converges for the short-term cross-correlated processes as shown in Proposition 2.5. The logic of the test is illustrated in Figure 4.3 where we compare the weakly and strongly short- and

long-range cross-correlated processes as in the previous section. It is evident that the differences in behavior of the covariances of partial sums are even more severe than for the previous test. Based on these observations, we propose a following testing statistic and test.

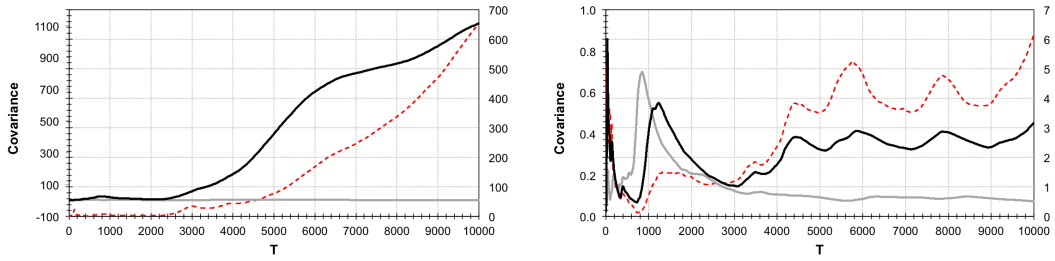


Figure 4.3: **Scaling of covariance of partial sums.** On the left chart, a comparison of strong short- and long-term cross-correlations with AR(1) process with $\theta = 0.9$ (grey) and ARFIMA(0,0.4,0) (black), both with perfectly correlated innovations (values on the left axis), with a ratio between the two in red (values on the right axis) is presented. On the right chart, a comparison of weak short- and long-term cross-correlations with AR(1) process with $\theta = 0.1$ (grey) and ARFIMA(0,0.1,0) (black), both with perfectly correlated innovations (values on the left axis) is presented. Ratio between the two is in red with values on the right axis.

Definition 4.2 (Partial sums covariance divergence test). Let processes $\{x_t\}$ and $\{y_t\}$, with $t = 1, \ldots, T$, be jointly wide-sense stationary processes with respective partial sums $\{X_t\}$ and $\{Y_t\}$. A partial sums covariance statistic γ_T is defined as

$$\gamma_T = \frac{\mathrm{Cov}(X_T, Y_T)}{T}. \tag{4.3}$$

The null and the alternative hypotheses are stated as "short-range cross-correlations" and "long-range cross-correlations", respectively. In the same manner as for the ACC test, we are always dealing with finite series and the actual value of the testing statistic would depend on the memory strength δ as well as the correlation level between series. Such a dependence is illustrated in Figure 4.4 where we observe, again similarly to the ACC test, that the mean value of the testing statistic increases with the memory strength as well as with the correlation level. The simulated processes have the same parameters as for the ACC test. However, compared to the ACC test, the difference between levels of the testing statistic under null and alternative hypotheses is of one order of magnitude, which is much more pronounced compared to the previous test. Variance of the testing statistics for both kinds of memory seem rather

independent of the level of correlations between innovations. However, the variance increase with both θ and d. The levels of standard deviation which are attained under the alternative hypothesis are as much as two orders of magnitude higher for the long-term memory case compared to the short-term one. Due to the dependence of the statistical properties of the γ_T statistic on the parameters, we again utilize the moving-block bootstrap procedure to obtain p-values for a two-sided test.

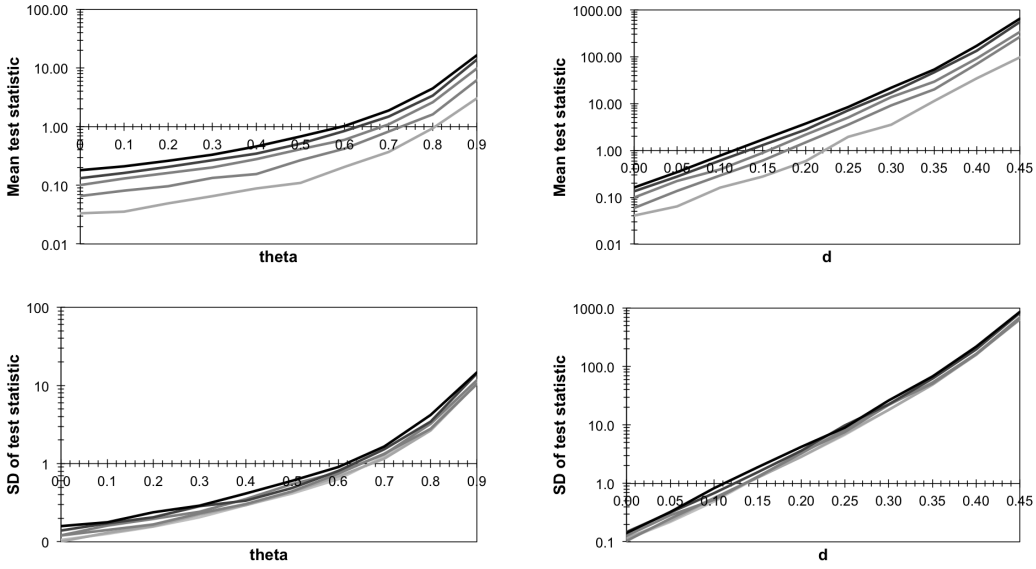


Figure 4.4: **Mean values and standard deviations of PSCD test.** Test statistic γ_{5000} for differently correlated processes. Correlation between innovations varies between 0 and 1 with a step of 0.2 and the darker the line in the chart is, the higher the correlation is. On the left, correlated AR(1) processes with θ ranging between 0 and 0.9 with a step of 0.1 are shown. On the right, correlated ARFIMA(0,d,0) processes with d ranging between 0 and 0.45 with a step of 0.05 are shown. Means are based on 1,000 simulations with a time series length of 5,000 and presented in a semi-log scale for better legibility.

Monte-Carlo-based sizes and powers of the PSCD test are summarized in Tables B.13 and B.14. We utilize the same set of processes as for the DCCA-based and ACC tests to control for various types and strengths of memory. Table B.13 shows the size of the test for correlated noise and short-range correlated processes. The size is practically equal to the ideal value for almost all combinations of memory, correlation between innovations and time series length. Even for the longest series with the strongest correlations, the size is equal to 0.008, 0.068, 0.113 for theoretical values of 0.01, 0.05 and 0.1, respectively.

Compared to the previous two tests, PSCD dominates both DCCA-based and ACC tests in its size under the null hypothesis of short-range cross-correlations.

For the long-range correlated processes, we observe relatively low values for almost all combinations of length, memory and correlations. Even for the series of 5,000 observations with $d = 0.4$ and $\rho = 0.9$, we have power of 0.519, 0.588 and 0.653 for the significance levels of 0.01, 0.05, and 0.1, respectively. Nevertheless, the values are still quite satisfactory and only show that there is a trade-off between size and power of the test.

4.4 Rescaled covariance test

Motivated by the works of Giraitis *et al.* (2003) and Lavancier *et al.* (2010), we propose a new test for the presence of long-range cross-correlations between two series. The test, which we call the rescaled covariance test, is based on non-summability of the cross-correlations (Proposition 2.1), on the scaling of the partial sums covariance (Proposition 2.3) and on the diverging limit of the covariance of the partial sum (Proposition 2.4). Before proposing the test itself, we need to define the heteroskedasticity and auto-correlation consistent (HAC) estimator of the cross-covariance $s_{xy,q}$ based on Giraitis *et al.* (2003) and Lavancier *et al.* (2010).

Definition 4.3 (HAC-estimator of covariance). Let processes $\{x_t\}$ and $\{y_t\}$ be jointly wide-sense stationary with a cross-covariance function $\gamma_{xy}(k)$ for lags $k \in \mathbb{Z}$. The heteroskedasticity and auto-correlation consistent estimator of $\gamma_{xy}(0)$ is defined as

$$\widehat{s_{xy,q}} = \sum_{k=-q}^{q} \left(1 - \frac{|k|}{q+1}\right) \widehat{\gamma_{xy}}(k), \tag{4.4}$$

where q is a number of lags of the cross-covariance function taken into consideration and the cross-covariances are weighted with Barlett-kernel weights.

The basic idea behind the rescaled covariance test (RCT) is to utilize the divergence of covariances of the partial sums of the long-range cross-correlated processes but also the convergence of the short-range cross-correlated processes and at the same time controlling for different levels of correlations in the case of the short-term memory utilizing $\widehat{s_{xy,q}}$. The rescaled covariance test is then defined as follows:

Definition 4.4 (Rescaled covariance test). Let processes $\{x_t\}$ and $\{y_t\}$, with $t = 1, 2, \ldots, T$, be jointly wide-sense stationary processes with a cross-covariance function $\gamma_{xy}(k)$ for $k \in \mathbb{Z}$ and with respective partial sums $\{X_t\}$ and $\{Y_t\}$. Assuming that $\sum_{k=-\infty}^{+\infty} \gamma_{xy}(k) \neq 0$, a rescaled covariance statistic $M_{xy,T}(q)$ is defined as

$$M_{xy,T}(q) = q^{\hat{d}_x + \hat{d}_y} \frac{\widehat{\mathrm{Cov}}(X_T, Y_T)}{T \widehat{s_{xy,q}}}, \tag{4.5}$$

where $\widehat{s_{xy,q}}$ is the HAC-estimator of the covariance between $\{x_t\}$ and $\{y_t\}$ (Definition 4.3), $\widehat{\mathrm{Cov}}(X_T, Y_T)$ is the estimated covariance between partial sums X_T and Y_T, and $\hat{d}_x$ and $\hat{d}_y$ are estimated fractional integration parameters for separate processes $\{x_t\}$ and $\{y_t\}$, respectively.

Similarly to the tests for long-range dependence in the univariate series which are based on the modified variance as in Equation 4.4, such as the rescaled variance (Giraitis *et al.* 2003) and the modified rescaled range analysis (Lo 1991), the choice of parameter q is crucial. If the parameter is too low, the strong short-range cross-correlations can be detected as the long-range cross-correlations and reversely, if the parameter is too high, the true long-range cross-correlations can be filtered out as the short-range ones. This issue is discussed later. Returning to the construction of RCT, the motivation was to construct a test which would have a test statistic that would be (at least partially) independent of the parameters included in the null hypothesis. In this case, the null and alternative hypotheses remain the same as for the previous two tests – the null hypothesis of short-range cross-correlated processes and the alternative of cross-persistent processes. Therefore, it is desirable to have a testing statistic independent of the correlation level of the short-range cross-correlated processes as well as of the memory parameter θ. In the same manner as for the previous two tests, we present Figure 4.5 where the means and standard deviations of the testing statistics are shown for both short- and long-term memory cases with varying parameters.

For short-range cross-correlated processes, we observe that the mean value is remarkably stable for parameters up to $\theta = 0.7$ regardless of the correlation between innovations. For higher values, the statistic deviates which can be, however, attributed to the fact that we applied $q = 30$ for estimation of the test statistic and that is evidently insufficient for such a strong memory. Interestingly, the mean value of the test statistic for $0 \leq \theta \leq 0.7$ practically overlays with the testing statistic of the rescaled variance test of Giraitis *et al.* (2003)

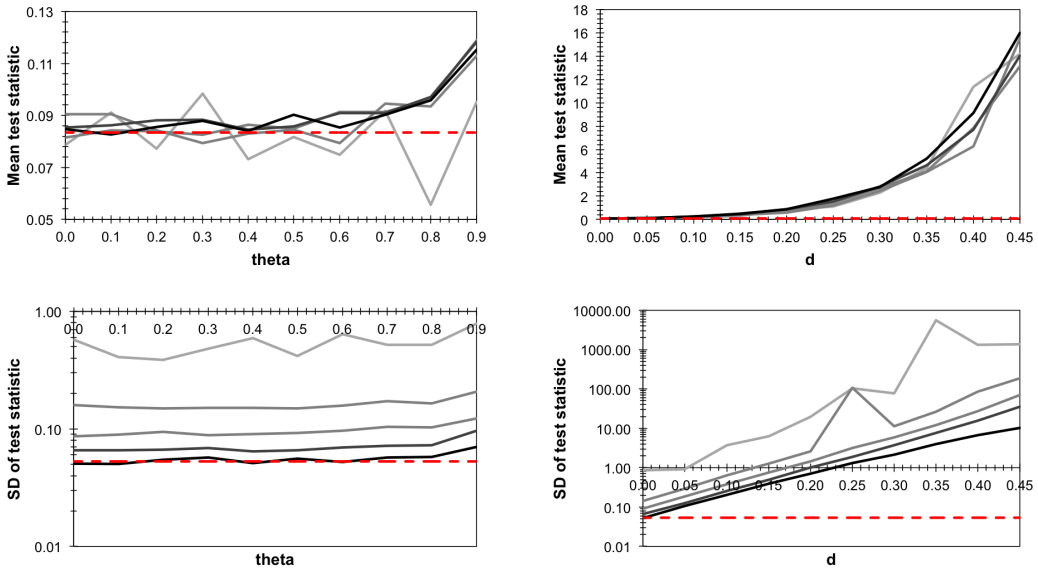


Figure 4.5: **Mean values and standard deviations of RCT test.** Test statistic $M_{xy,5000}(30)$ for differently correlated processes. Correlation between innovations varies between 0.2 and 1 with a step of 0.2 and the darker the line in the chart is, the higher the correlation is. On the left, correlated AR(1) processes with θ ranging between 0 and 0.9 with a step of 0.1 are shown. On the right, correlated ARFIMA(0,d,0) processes with d ranging between 0 and 0.45 with a step of 0.05 are shown. Means are based on 1,000 simulations with a time series length of 5,000 and presented in a semi-log scale for better legibility.

which is defined as

$$U = \int_0^1 \left(W_t^0\right)^2 dt - \left(\int_0^1 W_t^0 dt\right)^2 . \tag{4.6}$$

where W_t^0 is a standard Brownian bridge. Mean value of statistics U is equal to $1/12$, which is represented by a red line in Figure 4.5. In the figure, we also show behavior of the standard deviation of the statistic. Even though it is evidently dependent on the correlation between innovations of the AR(1) processes, it is remarkably stable across different levels of θ. Importantly, the variance decreases with increasing correlation between innovations which is a very desirable property not observed for the previous two tests. For the perfectly correlated innovations of the series, the standard deviation of the statistics even attains the levels for U – $1/\sqrt{360}$. For the long-range cross-correlated processes, we observe that the mean value of the statistic increases with d as expected. Again, the mean value is very stable with respect to the correlation of innovations. However, the variance of the estimator increases with d parameter and is also dependent on the correlations between innovations. Even though the $M_{xy,T}(q)$ statistic shows some very desirable properties, we still opt to base our decision in favor or against the alternative hypothesis based on the MBB procedure, mainly due to dependence of the variance of the estimator on the correlations level.

To examine the size and power of the test, we use the same setup as in the previous cases. For correlated but not cross-correlated processes (Table B.15), we observe that the test is more precise with increasing correlation ρ between innovations of the processes. For $\rho = 0.9$, the size of the test practically matches the set significance levels. The size of the test gets better with increasing q and practically does not vary with time series length T. Even for $\rho = 0$, the test shows very reasonable size despite no correlations between innovations violating the assumption of the test that $\sum_{k=-\infty}^{+\infty} \gamma_{xy}(k) \neq 0$. Practically the same results are observed for the short-range cross-correlated processes as shown in Table B.16. The sizes practically overlay with the theoretical values of the significance levels. These are very strong results in favor of the rescaled covariance test as it is practically intact by even very strong short-term memory. The combination of the moving-block bootstrap and HAC-estimator of covariance is evidently able to sufficiently control for possible short-term memory biases in case of the RCT test.

For long-range cross-correlated processes, we compare cases when $d_x = d_y =$

0.1 and $d_x = d_y = 0.4$ to distinguish between weak and strong cross-persistence. We assume these values of d_x and d_y in the testing procedure. The power of the test is relatively low for the weak cross-persistence case (Table B.19). We, however, observe several interesting points. First, the power of the test is very similar regardless the correlation level between disturbances (leaving the case of uncorrelated disturbances aside as it violates the assumptions of the test). Second, the power of the test increases with the time series length. Third, the power increases rapidly with increasing α. And fourthly, the power of the test even increases with an increasing q, which is caused by the $q^{d_x+d_y}$ factor in the testing statistic which well compensates for high q. For the strong cross-persistence (Table B.20), the power of the test increases considerably and the four features of the test are the same as in the previous case. As expected, the test is more powerful with increasing ρ, i.e. the cross-persistence is more stable. The power of the test increases to as high as 0.967 for some cases. The test thus shows very good statistical characteristics and is well able to distinguish between short-range and long-range cross-correlations.

4.5 Brief overview

We have analyzed four tests for detecting long-range cross-correlations, three of which are newly introduced in this thesis. The new tests strongly dominate the detrended cross-correlation coefficient test of Zebende (2011) and Podobnik *et al.* (2011) in terms of power and size of the test. Specifically, the test detects the long-range cross-correlations almost every time the series are at least somehow correlated, regardless the type of memory. The power of the test is thus very high, which, however, comes with completely inapplicable levels of size of the test. This could be partially overcome by applying the moving-block bootstrap mechanism in the hypothesis testing as is done in tests we propose. However, the test as it is is already computationally demanding and under the bootstrapping, it might become unbearable and very time-consuming. On the brighter side, the idea behind the test can be utilized to develop similar tests in the HXA and DMCA frameworks both of which would be more appropriate for the moving-block bootstrap framework as these two methodologies do not demand box-splitting.

Out of the three tests we newly propose – the aggregate cross-correlations test, the partial sums covariance divergence test and the rescaled covariance test – the lastly proposed one gives the most balanced combination of power

and size, followed by the aggregate cross-correlations test. The partial sums covariance divergence test is rather strongly dominated by the other two. All three tests can be easily compared as the hypothesis rejection decision is based on the moving-block bootstrap p-value for each of them. Interestingly, the specification of the three tests combined with the application of the moving-block bootstrapping is able to control for even very strong short-range dependence and the tests are almost untouched by such bias.

We suggest to use a combination of the tests in any applied work on long-range cross-correlations to provide statistically more robust results. Based on the Monte-Carlo-simulations-based sizes and powers of the tests, we propose a combination of the rescaled covariance and aggregate cross-correlations tests. Such testing should be conducted as an initial step before any bivariate Hurst exponent or power law coherency estimations.

Chapter 5

Estimators of bivariate Hurst exponent

In this chapter, we focus on the estimators of the bivariate Hurst exponent. Three time domain – heuristic – methods are introduced as well as three frequency domain estimators. We present the detrended cross-correlation analysis, which was proposed by Podobnik & Stanley (2008) and since then, it has been applied in various fields such as hydrology (Hajian & Sadegh Movahed 2010), seismology (Shadkhoo & Jafari 2009), agriculture (He & Chen 2011b), finance (Podobnik *et al.* 2009b; Cao *et al.* 2012), traffic (Zhao *et al.* 2011) and other fields. We then introduce the height cross-correlation analysis, which has been introduced by the author of this thesis (Kristoufek 2011). In the same year, He & Chen (2011a) introduced the detrending moving-average cross-correlation analysis, which is presented as well. In the frequency domain, we discuss the averaged periodogram estimator of Sela & Hurvich (2012) and we further develop two new methods for the estimation of the bivariate Hurst exponent – cross-periodogram method and local X-Whittle estimator. For each of the estimators, we compare their finite sample bias, variance and mean squared errors for processes from Chapter 3. We deal only with the long-range cross-correlated processes as we have shown in the previous chapter that we are well able to test whether the analyzed series are or are not long-range cross-correlated. Apart from the estimation of the bivariate Hurst exponent H_{xy}, we also present the estimators of the power law coherency H_ρ introduced in Subsection 2.2.3. In order to do so, we introduce estimators based on the same ideas as the estimators of the bivariate Hurst exponent. We stress here that for all estimators, the underlying processes are assumed to be long-range cross-correlated according

to Definition 2.4 and no further assumptions are stated.

The finite sample properties of estimators are analyzed from various viewpoints. First, we check whether the estimators are able to find the correct H_{xy} when it holds that $H_{xy} = \frac{H_x+H_y}{2}$. This is done for ARFIMA processes with correlated innovations for two cases – $H_x = H_y = 0.6$ and $H_x = H_y = 0.9$ – to control for weak and strong long-range cross-correlations. Second, we are interested in a situation when one of the processes is short-range dependent as short-range correlations are frequently discussed even for the univariate case. In order to do so, we analyze the case when we have an ARFIMA and an AR(1) process with correlated innovations, fixing the long memory parameter to $H = 0.9$ and by varying the short memory parameter $\theta = 0.1, 0.5, 0.8$, we discuss the effect of a strengthening short memory. Third, we inspect how the estimators perform under the power law coherency. To check this, we simulate the series from the mixed-correlated ARFIMA processes with $H_1 = H_4 = 0.9$ and $H_2 = H_3 = 0.7$ while only the innovations ε_2 and ε_3 are correlated. Fourth, we use the same setting as in the previous case but we are interested in the power law coherency itself rather than the bivariate Hurst exponent. In all cases, we analyze three cases of the time series length – $T = 500, 1000, 5000$ – and we are interested in the effect of the strength of the correlation between innovations so that we check the finite sample properties for the correlation levels 0.1, 0.5 and 0.9. Other parameters of the simulations are specified for each estimator later if necessary. For all estimators, we are interested in their bias, standard error and mean squared error.

5.1 Time domain estimators

Time domain estimators are based on Proposition 2.3 which states that the covariance of the partial sums of long-range cross-correlated processes scales with the time series length following a power law. The estimators differ by different measures of covariation and detrending applied in the procedures similarly to the univariate case and estimators such as the rescaled range analysis (Hurst 1951), detrended fluctuation analysis (Peng *et al.* 1993; 1994; Kantelhardt *et al.* 2002), variance method (Kokoszka & Taqqu 1996), absolute value method (Taqqu & Teverovsky 1996), generalized Hurst exponent approach (Di Matteo *et al.* 2003), height-height correlation analysis (Barabasi *et al.* 1991; Barabasi & Vicsek 1991; Alvarez-Ramirez *et al.* 2002), and other methods.

A serious issue arises in the generalization of the time domain univariate Hurst exponent estimators for the bivariate setting. The univariate estimators are all based on some version of variance for differently filtered or differenced series. However, when we generalize the estimators, we can either use a covariance measure between two processes or the absolute value of the covariance measure. The former possibility is closer to the idea of long-range cross-correlations since Proposition 2.3 is based on covariances and not absolute value of covariances. However, it might happen that the estimated covariance for a specific scale is in fact negative so that it cannot enter the power-law scaling estimation and thus increases the variance of the estimator. The latter case does not enable the value to be negative and thus keeps the variance of the estimator lower. However, the estimate of the bivariate Hurst exponent might be dragged to the average value of the separate Hurst exponent even if it is not equal to it. This issue has been shortly discussed in Kristoufek (2011) through the definition of multifractal processes. To capture and understand the issue in more detail, we analyze both versions for each of the three time domain estimators.

5.1.1 Detrended cross-correlation analysis

Detrended cross-correlation analysis (DCCA, or DXA) is the most frequently used method for the estimation of the bivariate Hurst exponent in the time domain. Podobnik & Stanley (2008) construct the method as a bivariate generalization of the detrended fluctuation analysis (DFA), which is again probably the most popular heuristic method of estimating the (generalized) Hurst exponent (Peng *et al.* 1993; 1994; Kantelhardt *et al.* 2002). DCCA was further generalized for the multifractal analysis by Zhou (2008) and the multifractal detrended cross-correlation analysis (MF-DXA) was developed. Jiang & Zhou (2011) altered the filtering procedure in MF-DXA with using the moving averages to create the multifractal detrending moving average cross-correlation analysis (MF-X-DMA). DCCA was also used to construct statistical tests for the presence of long-range cross-correlations between two series (Zebende 2011; Podobnik *et al.* 2011). These tests are discussed in more detail in Chapter 4.

In the DCCA procedure, we consider two long-range cross-correlated series $\{x_t\}$ and $\{y_t\}$ with $t = 1, \ldots, T$. Their respective integrated processes $\{X_t\}$ and $\{Y_t\}$, defined as $X_t = \sum_{i=1}^{t} x_i - \bar{x}$ and $Y_t = \sum_{i=1}^{t} y_i - \bar{y}$, for $t = 1, \ldots, T$, are divided into overlapping boxes of length s so that $T - s + 1$ boxes are constructed. In each box between j and $j + s - 1$, the linear fit of a time trend

is constructed so that we get $\widehat{X_{k,j}}$ and $\widehat{Y_{k,j}}$ for $j \leq k \leq j+s-1$. The covariance between residuals in each box is defined as

$$f^2_{DCCA}(s,j) = \frac{\sum_{k=j}^{j+s-1}(X_k - \widehat{X_{k,j}})(Y_k - \widehat{Y_{k,j}})}{s-1}. \tag{5.1}$$

The covariances are finally averaged over the blocks of the same lengths s and we obtain the detrended covariance as

$$F^2_{DCCA}(s) = \frac{\sum_{j=1}^{T-s+1} f^2_{DCCA}(s,j)}{T-s}. \tag{5.2}$$

For the long-range cross-correlated processes, the covariance scales as

$$F^2_{DCCA}(s) \propto s^{2H_{xy}}. \tag{5.3}$$

The estimate of the bivariate Hurst exponent is obtained by the log-log regression on Equation 5.3. Similarly to DFA and MF-DFA, there are several ways of how to treat overlapping and non-overlapping boxes of length s, compare e.g. Peng *et al.* (1993), Taqqu *et al.* (1995), Kantelhardt *et al.* (2002), Barunik & Kristoufek (2010) and Kristoufek (2010c). In the following text, we use non-overlapping boxes with a step between s equal to 10 due to computational efficiency. Additionally, we also consider the case when $f^2_{DCCA}(s,j)$ in Equation 5.1 is treated as an absolute value.

5.1.2 Height cross-correlation analysis

Kristoufek (2011) introduces the multifractal height cross-correlation analysis (MF-HXA) as a bivariate generalization of the height-height correlation analysis (Barabasi *et al.* 1991; Barabasi & Vicsek 1991; Alvarez-Ramirez *et al.* 2002) and the generalized Hurst exponent approach (Di Matteo *et al.* 2003; 2005; Di Matteo 2007), which are often labeled simply as HHCA and GHE, respectively.

MF-HXA is constructed to analyze the multifractal properties of bivariate series similarly to MF-DXA. We generalize the q-th order height-height correlation function for two simultaneously recorded series. Let us consider two integrated series $\{X_t\}$ and $\{Y_t\}$ with time resolution ν and $t = \nu, 2\nu, ..., \nu\lfloor\frac{T}{\nu}\rfloor$, where $\lfloor\rfloor$ is a lower integer sign. For better legibility, we denote $T^* = \nu\lfloor\frac{T}{\nu}\rfloor$, which varies with ν, and we write the τ-order difference as $\Delta_\tau X_t \equiv X_{t+\tau} - X_t$ and $\Delta_\tau X_t Y_t \equiv \Delta_\tau X_t \Delta_\tau Y_t$. Height-height covariance function is then defined as

$$K_{xy,q}(\tau) = \frac{\nu}{T^*} \sum_{t=1}^{T^*/\nu} |\Delta_\tau X_t Y_t|^{\frac{q}{2}} \equiv \langle |\Delta_\tau X_t Y_t|^{\frac{q}{2}} \rangle \tag{5.4}$$

where time interval τ generally ranges between $\nu = \tau_{min}, \ldots, \tau_{max}$. Scaling relationship between $K_{xy,q}(\tau)$ and the generalized bivariate Hurst exponent $H_{xy}(q)$ is obtained as

$$K_{xy,q}(\tau) \propto \tau^{qH_{xy}(q)}. \tag{5.5}$$

Obviously, MF-HXA reduces to the height-height correlation analysis of Barabasi *et al.* (1991) and Barabasi & Vicsek (1991) for $\{X_t\} = \{Y_t\}$ for all $t = 1, \ldots, T$. Note that it makes sense to analyze the scaling according to Equation 5.5 only for detrended series $\{X_t\}$ and $\{Y_t\}$ and only for $q > 0$ (Di Matteo 2007). A type of detrending can generally take various forms – polynomial, moving averages and other filtering methods – and is applied for each time resolution ν separately.

For the analysis of long-range cross-correlations, it suffices to consider only the case $q = 2$ as for MF-DXA and DCCA case, MF-HXA then reduces to height cross-correlation analysis – HXA – and we write $H_{xy} \equiv H_{xy}(2)$. For this specific case, we don't take the square root in Equation 5.4 since $\frac{q}{2} = 1$ which means that the absolute values are not needed. Asymptotically, the variation makes no difference (Zhou 2008) but it might matter for finite samples. Therefore, we consider two versions of the height-height covariance function in the simulations study – the one with covariance and the one with covariance of absolute values. As for now, we stick with the original version of Kristoufek (2011). Height-height covariance function $K_{xy,2}(\tau)$ is defined as

$$K_{xy,2}(\tau) = \frac{\nu}{T^*} \sum_{t=1}^{T^*/\nu} |\Delta_\tau X_t Y_t| \equiv \langle |\Delta_\tau X_t Y_t| \rangle \tag{5.6}$$

For processes $\{X_t\}$ and $\{Y_t\}$ with long-range cross-correlated increments $\{x_t\}$ and $\{y_t\}$, we expect that the height-height covariance function scales as

$$K_{xy,2}(\tau) \propto \tau^{2H_{xy}}. \tag{5.7}$$

The estimated bivariate Hurst exponent is again obtained via the log-log regression. It has been argued that the best estimates and the most regular scaling is obtained for $\tau/T \to 0$ (Barabasi *et al.* 1991; Barabasi & Vicsek 1991)

and in the applications to finance and economics, it has been shown that the most appropriate setting is to use a fixed $\tau_{min} = 1$ and several values of τ_{max}, usually between 5 and 19 (or 20), and take the average Hurst exponent of these estimates as the best fit to the actual value, which practically means obtaining the jackknife estimate of Hurst exponent, see Di Matteo *et al.* (2003; 2005), Di Matteo (2007), Barunik & Kristoufek (2010), Kristoufek (2010b), and Barunik *et al.* (2012) for details.

5.1.3 Detrended moving-average cross-correlation analysis

Detrending moving average (DMA) was proposed as a method for estimating Hurst exponent by Alessio *et al.* (2002) motivated by the work of Vandewalle & Ausloos (1998). Even though the method is not directly connected to the power-law decay of auto-correlations nor to the scaling of variances of the partial sums nor the diverging power spectrum, it has been frequently applied mainly due to its computational efficiency. The connection between the estimator itself and the actual long-range dependence – that the variance of integrated series of the long-range dependent process follows a power-law with respect to the length of the moving window – has been shown numerically (Grech & Mazur 2005; Barunik & Kristoufek 2010).

To check whether the relationship holds also for the scaling of covariances, He & Chen (2011a) proposed a new method called detrended moving-average cross-correlation analysis (DMCA). Note that this should not be confused with the MF-X-DMA method of Jiang & Zhou (2011), which only applies the moving average filtering to the DCCA or MF-DXA methodology or with 2D-DMA of Carbone (2007) or the method of Arianos & Carbone (2009).

For two series $\{x_t\}$ and $\{y_t\}$ and their respective partial sums $\{X_t\}$ and $\{Y_t\}$, the detrended covariance $F^2_{DMCA}(\kappa)$ is defined as

$$F^2_{DMCA}(\kappa) = \frac{1}{T-\kappa+1} \sum_{i=\lfloor \kappa/2 \rfloor + 1}^{T-\lfloor \kappa/2 \rfloor} \Big(X_i - \widetilde{X_i(\kappa)} \Big) \Big(Y_i - \widetilde{Y_i(\kappa)} \Big), \tag{5.8}$$

where $\widetilde{X_i(\kappa)}$ and $\widetilde{Y_i(\kappa)}$ are respective non-weighted centered moving averages at time point i with a moving average window of length $\kappa = 1, 3, 5, \ldots, \kappa_{max}$, where κ_{max} is an odd integer. The form of the moving average can take various forms (centered, backward, forward, weighted or unweighted). For the long-

range cross-correlated processes $\{x_t\}$ and $\{y_t\}$, we expect to observe

$$F^2_{DMCA}(\kappa) \propto \kappa^{2H_{xy}}. \tag{5.9}$$

Note that compared to DCCA, the series is not split into boxes which makes the method much more straightforward and computationally efficient.

5.1.4 Power law coherency estimation

The power law coherency is defined in Chapter 2 as $H_\rho = H_{xy} - \frac{H_x+H_y}{2} \leq 0$. Even though the basic idea for the power law coherency lays in the behavior of the squared spectrum coherence near the origin, it can be easily translated into the behavior of the covariance measures for high scales. Using the definition of the power law coherency, we can simply use the covariance measures from DCCA, HXA and DMCA to construct the parallel to the squared coherency as a squared correlation coefficient as it holds that

$$\rho^2_{DCCA} = \frac{|F^2_{DCCA}(s)|^2}{F^2_{DFA,x}(s)F^2_{DFA,y}(s)} \propto \frac{s^{4H_{xy}}}{s^{2H_x}s^{2H_y}} = s^{4(H_{xy}-\frac{H_x+H_y}{2})} = s^{4H_\rho} \tag{5.10}$$

$$\rho^2_{HXA} = \frac{|K^*_{xy,2}(\tau)|^2}{K^*_{x,2}(\tau)K^\tau_{y,2}(\tau)} \propto \frac{\tau^{4H_{xy}}}{\tau^{2H_x}\tau^{2H_y}} = \tau^{4(H_{xy}-\frac{H_x+H_y}{2})} = \tau^{4H_\rho} \tag{5.11}$$

$$\rho^2_{DMCA} = \frac{|F^2_{DMCA}(\kappa)|^2}{F^2_{DMA,x}(\kappa)F^2_{DMA,y}(\kappa)} \propto \frac{\kappa^{4H_{xy}}}{\kappa^{2H_x}\kappa^{2H_y}} = \kappa^{4(H_{xy}-\frac{H_x+H_y}{2})} = \kappa^{4H_\rho} \tag{5.12}$$

Note that $K^*_{xy,2}(\tau)$ is defined without the absolute values:

$$K^*_{xy,2}(\tau) = \frac{\nu}{T^*}\sum_{t=1}^{T^*/\nu} \Delta_\tau X_t Y_t \equiv \langle \Delta_\tau X_t Y_t \rangle \tag{5.13}$$

The power law coherency is estimated through a log-log regression on Equation 5.10-5.12 obtaining $\widehat{H_{\rho,DCCA}}$, $\widehat{H_{\rho,HXA}}$ and $\widehat{H_{\rho,DMCA}}$, respectively. The fact that the covariances in the definitions of DCCA and DMCA are in their original form and the one in HXA is defined without the absolute value should ensure a better fit to the true bivariate Hurst exponent. The fact that the covariance enters the definition of squared correlations in a squared form ensures that the power-law scaling can be easily found and estimated.

5.1.5 Finite sample properties

Statistical performance of detrended fluctuation analysis has been shown to be dependent on a choice of s_{min} and s_{max}, i.e. the minimum and the maximum scales taken into consideration (Weron 2002; Grech & Mazur 2005; Kristoufek 2010c; Barunik & Kristoufek 2010). In the same way, DCCA can perform differently for various settings. We set $s_{max} = T/5$, which is standardly done in the literature and we manipulate $s_{min} = 10, 20, 50$ for $T = 500$ and $s_{min} = 10, 50, 100$ for the other two cases, $T = 1000, 5000$. The results of the simulations for DCCA are summarized in Table C.1-C.7.

For ARFIMA processes with correlated innovations, there are several interesting findings (Table C.1-C.2). First, both the original and the absolute value based definitions of DCCA are biased downwards, while the absolute value based version is less biased than the original one. Second, standard deviation of the estimators increases with the increasing s_{min}. Third, for the short series of $T = 500$, the standard deviations of the original DCCA are much higher than for the absolute value version. Fourth, the original version of DCCA is practically useless for weakly correlated processes as for the case when the correlation between innovations is only 0.1, there are so many cases when the log-log regression could not be performed that it does not make sense to report the results. We shall see that this is true for all the time domain methods. Fifth, for the absolute value version of the estimator, the standard deviation increases with the correlation between innovations, which is probably caused by the fact that the absolute value version of the estimator is pushed towards the average of the separate Hurst exponents and the correlation between innovations only disturbs this effect. Sixth, the inverse is true for the original version of DCCA, i.e. the variance of the estimator decreases with the correlation between innovations. Seventh, bias and variance of the estimator decreases with time series length. And eighth, mean squared error decreases with the time series length.

For long-range cross-correlations arising from the combination of ARFIMA and AR(1) processes, the results are summarized in Table C.3-C.5. In general, both DCCA-based estimators are quite robust to the presence of short-range dependence even for a very strong memory of $\theta = 0.8$. More specific findings follow. First, for the weak short-term memory ($\theta = 0.1$), both estimators are still biased downwards, which is more profound for the original DCCA procedure. Interestingly, the mean squared error decreases with increasing s_{min}

for both estimators which is rather unexpected as the short-range dependence usually mostly affects the lower scales. Variance of the original estimator is higher than for the absolute value version of the estimator. Again, the original DCCA estimator performs very poorly for the lowest values of correlation between innovations and it is not reported for any of the processes. Second, for the processes with more profound short memory, we observe an expected behavior of bias – it reduces with the increasing minimum scale s_{min}. However, the variance of the estimator increases with s_{min} so that it more than offsets the bias gains and the total mean squared error increases with the increasing s_{min} for both versions of the estimator. In terms of a choice of the estimator and parameter setting, we would suggest to use low minimum scale for both versions of the estimator as it provides a good balance between bias and variance. Note that for $T = 5000$, $\theta = 0.8$ and the strongest correlation between innovations, the mean squared errors of the estimator equal to 0.0033 and 0.0036 for $s_{min} = 10$ for the original version and the absolute value version of the estimator, respectively, and does not increase markedly for the other minimum scales, which strongly dominates even the frequency domain estimators introduced later in the text. The DCCA estimators are thus very robust to the potential short-term memory bias.

The situation changes for the mixed-correlated ARFIMA processes (Table C.6). Disturbingly, the estimates based on the original version of DCCA are not able to estimate H_{xy} very frequently even for the correlation between innovations of 0.5, due to the number of cases when the detrended variances are negative for several scales. We thus provide the results only for the absolute value version of DCCA. First, the bias increases with the time series length and the estimator is thus not consistent. Second, the bias and variance, and thus also mean squared error, of the estimator increase with increasing s_{min}. Third, the bias decreases with the strength of correlations between innovations. Unfortunately, this decrease is only mild. The estimator is thus not a very good tool to estimate H_{xy} if it is not equal to the average of the separate Hurst exponents.

As an estimator of the power law coherency, the DCCA-based method does not attain desirable properties as shown in Table C.7. For low correlation between innovations, the bias reaches values above 0.4 and the situation does not improve much for the higher correlations mainly due to very high variance of the estimator. Even for the best case, which is the shortest series with $s_{min} = 10$ and correlation between innovations equal to 0.9, we have relatively

low bias of 0.03 but standard deviation of the estimator is still equal to approximately 0.2. Nonetheless, the bias and variance decrease considerably with the strength of correlation between innovations. Both bias and variance increase with the increasing minimum scale s_{min}. We cannot generally say that the bias and variance decrease with the time series as these results vary for different combinations of $\rho_{\varepsilon\nu}$ and s_{min}. All in all, the DCCA-based estimator of H_ρ does not give satisfying results.

For height cross-correlation analysis, we are mainly interested in its performance using various maximum scales τ_{max}. As a starting point, we use $\tau^*_{max} = 20$, which is frequently suggested in the literature (Di Matteo *et al.* 2003; 2005; Di Matteo 2007; Kristoufek 2010b; 2011). Moreover, we check the maximum scales of 50 and 100. To obtain more stable estimates of the bivariate Hurst exponent, we apply the bootstrapping procedure which estimates the Hurst exponent as a mean of Hurst exponents estimated using $\tau_{min} = 1$ and $\tau_{max} = 5, \ldots, \tau^*_{max}$. The simulated processes are the same as for DCCA methods and the results are summarized in Table C.8-C.14.

For correlated ARFIMA processes (Table C.8-C.9), we observe that both versions of the estimator are downward biased and the bias is more severe for stronger long-range cross-correlations. For lower levels of correlations, i.e. the lower correlations of innovations, the original (absolute value based) approach strongly outperforms the adjusted version. However, the differences become negligible for the innovations correlation of 0.9. In general, it holds that the bias is more severe and the variance of the estimators increase with increasing τ^*_{max}. Similarly to the DCCA estimators, we find that the behavior of variance of the estimators differs for the two approaches. For the absolute value based HXA, the variance slightly increases with the correlation of innovations while the opposite is true for the alternative specification, which is indeed more desirable and intuitive. The interpretation is the same as for DCCA – the absolute value approach to the time domain approaches pushes the estimates of the bivariate Hurst exponent to the average of the separate Hurst exponents, which is obviously not desirable. Nonetheless, the bias and variance of both versions of the estimator decrease with the time series length.

The HXA method is quite robust to the short memory bias as is shown for the combination of ARFIMA and AR(1) processes in Table C.10-C.12. The properties of the estimators depend on the strength of the short memory component. For weak short memory ($\theta = 0.1$), the properties are practically the same as for the previous cases – bias and variance increase with τ_{max} but de-

crease with the time series length with the same differences between the original and alternative definitions of the method. For the stronger short memory, we observe that even though the variance of the estimator still increases with τ_{max}, the bias decreases. Such disproportion is the most obvious for the longest time series length analyzed for which the mean squared error decreases with the increasing τ_{max}. Interestingly, the adjusted HXA outperforms the original HXA in a sense of the mean squared error for the longest series length and the strongest correlation between innovations. For the strong short memory with $\theta = 0.8$, both estimators are strongly upwards biased and the bias increases with the time series length for both specifications, making the estimator inconsistent. The variance, however, decreases with the time series length so that the mean squared error remains quite stable with varying length of the series. The adjusted version of HXA outperforms the original absolute value based version in bias but variance-wise, it is the other way around. In terms of the mean squared error, the adjusted version outperforms the original one for this level of short term memory.

Table C.13 summarizes the results for estimating the bivariate Hurst exponent for the mixed-correlated ARFIMA processes. We again observe, similarly to the DCCA case, that the estimator is upward biased. The bias increases with the time series length making the estimator inconsistent. Expectedly, the bias decreases with an increasing correlation between innovations. However, the decrease is rather mild. Connection of the bias to the maximum scale used varies for different time series lengths. The variance of the estimator decreases with time series length, increases with the maximum scale and remains practically unchanged for different levels of correlation between innovations. HXA estimator quite easily outperforms the DCCA approach for this type of processes. Note that we again report only the estimates for the original procedure as the method without the absolute values practically collapses in the same manner as for DCCA.

Performance of the HXA methodology for the power law coherency estimation is illustrated in Table C.14. Even though the estimator practically collapses for the low levels of correlation between innovations in a similar manner as the DCCA-based estimator, the situation improves rapidly for the higher correlation levels. Very importantly, the bias of the estimator decreases with the time series length, which has not been observed for DCCA, and the estimator thus seems to be consistent. Note, however, that such a statement cannot be supported with asymptotic theory and a rate of convergence cannot be eas-

ily commented on without further simulations. Nonetheless, this is a huge improvement over the DCCA-based estimator. The variance of the estimator decreases with the time series length as well, supporting the previous finding. Increasing the maximum scale increases the bias but decreases the variance of the estimator in most cases. However, the decrease in variance does not offset the increasing bias as the mean squared error increases with the maximum scale (apart from the least correlated cases).

The last time domain estimator we analyze is DMCA. The results of the Monte Carlo simulations are summarized in Table C.15-C.21. For weakly long-range cross-correlated processes based on correlated ARFIMA processes (Table C.15), the estimator is unbiased regardless the time series length or other parameters. The variance of the estimator increases with the length of the applied moving average and decreases with the time series length for both specifications of the DMCA method. The variance is in general lower for the absolute value based estimator. Again, we do not report the results for weakly correlated ($\rho = 0.1$) innovations for the original method as it provides only few actual estimates. For strongly long-range cross-correlated series (Table C.16), the estimator is downward biased. The bias and variance decrease with the time series length for both specifications of the estimator hinting consistency. However, the rate of convergence seems rather slow as the decrease of bias is rather weak among the analyzed time series lengths. The bias also decreases with κ_{max}. Reversely, the variance increases with κ_{max} and the total effect measured by the mean squared error varies with the time series length.

For the processes based on ARFIMA and AR(1), we observe that the short memory bias is not severe for weak short memory (Table C.17) while there is slight downward bias for the original version of the estimator. The variance of the estimator increases with the moving average window size as does the mean squared error. The bias, variance and mean squared error decreases with time series length for all cases. For the medium and strong short memory (Table C.18-C.19), the bias increases considerably and it becomes much higher than for the DCCA and HXA methods. Quite expectedly, the bias decreases with the maximum moving average window size for both medium ($\theta = 0.5$) and strong memory ($\theta = 0.8$). This is due to a stronger effect of short memory on the lower scales and thus lower moving average window sizes. In the same manner, the variance of the estimator increases with κ_{max}. The total effect results in the decreasing mean squared error with the maximum moving average size. Importantly, the DMCA estimates do not show a decreasing tendency of

the bias with the increasing time series length.

For the mixed-correlated ARFIMA processes (Table C.20), we observe that the bias and variance decrease with the time series length but increase with κ_{max}. The bias also decreases with increasing correlation between innovations while the variance increases slightly. DMCA strongly outperforms both DCCA and HXA in this aspect. This is also reflected in the performance of DMCA for estimating the power law coherency (Table C.21). Even though the finite sample properties are quite unstable across the various specifications, the estimator brings more favorable results than DCCA and HXA.

5.1.6 Comparison

Apart from the wide presentation of the finite sample properties of the time domain estimators, we can also focus on several specific underlying processes and compare the performance of the estimators. To see how the estimators behave with varying time series length T for the most interesting cases of possible short-term memory bias and power law coherency, we present Figure 5.1-5.3.

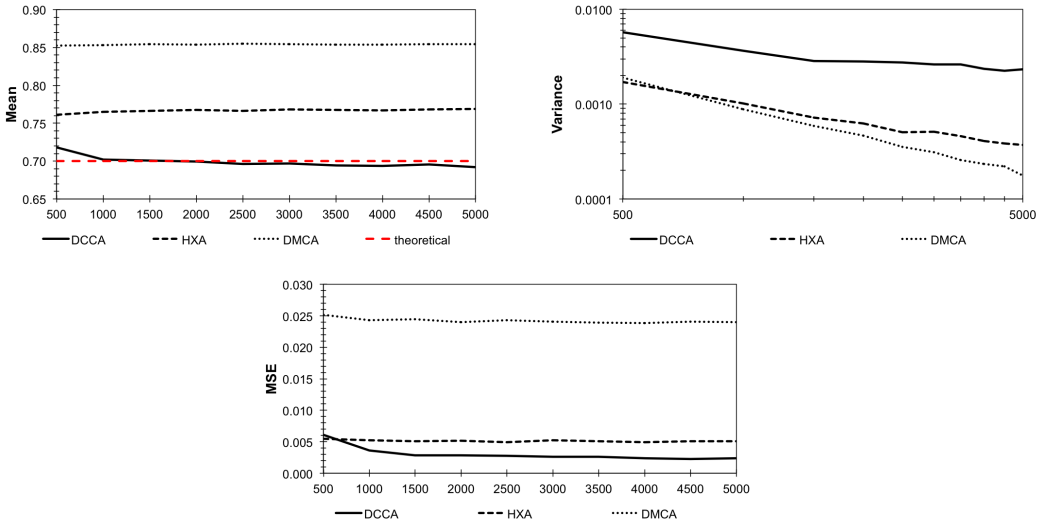


Figure 5.1: **Comparison of DCCA, HXA and DMCA estimators I.** Values are based on 1,000 simulations of ARFIMA(0,d,0) and AR(1) processes with perfectly correlated innovations, and $d = 0.4$ and $\theta = 0.5$. Time series length (x-axis) varies between 500 and 5,000 with a step of 500. Red line represents the true value of $H_{xy} = 0.7$.

In Figure 5.1, we compare the estimators for ARFIMA(0,d,0) and AR(1) processes with perfectly correlated innovations, $d = 0.4$ and $\theta = 0.5$ and changing time series length from 500 to 5,000 with a step of 500. For DCCA, we set

$s_{min} = 10$ and $s_{max} = T/5$; for HXA, we use the jackknife specification and $\tau_{min} = 1$ and $\tau_{max} = 5, \ldots, 20$; and for DMCA, we use $\kappa_{min} = 3$ and $\kappa_{max} = 21$. For all three estimators, we use the absolute value based versions due to their superiority in the power law coherency cases as discussed in the previous section. DCCA evidently dominates the other two estimators in its unbiasedness regardless the time series length. DMCA is the most biased estimator of the three for this specification. For all estimators, the bias is practically unaffected by the time series length. HXA and DMCA thus seem inconsistent for medium short-term memory bias. Even though DCCA evidently dominates the other two estimators in bias, it is the other way around for variance. For all estimators, the variance decreases with time series. For HXA and DMCA, the variance decays approximately following a power law. The variance of DCCA, however, decreases more slowly than a power law. DMCA shows the lowest variance apart from the very short series where it is dominated by HXA. Putting bias and variance together, mean squared error is very similar for DCCA and HXA regardless the time series length (DCCA shows slightly lower values), and these two strongly surpass DMCA.

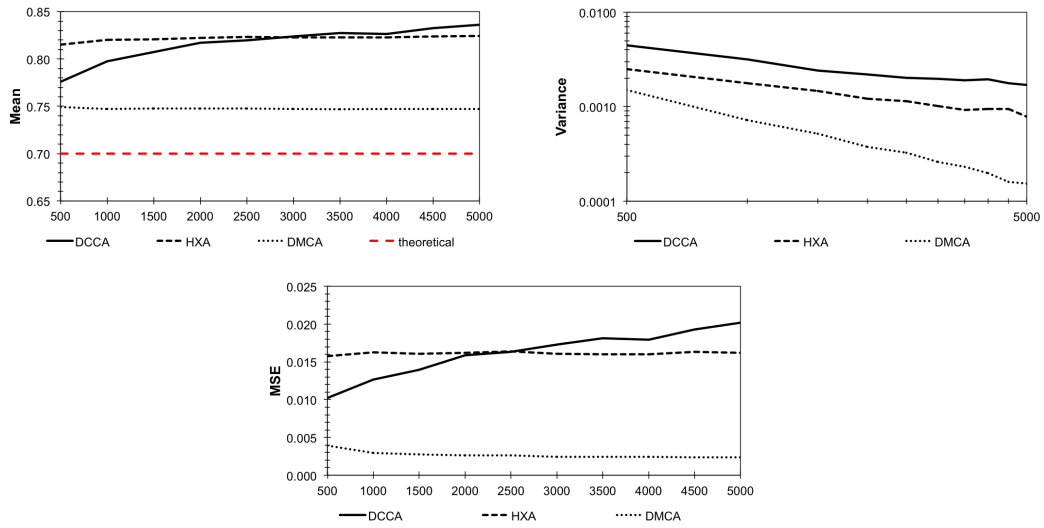


Figure 5.2: **Comparison of DCCA, HXA and DMCA estimators II.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$, $d_2 = d_3$ and perfectly correlated innovations ε_2 and ε_3. Time series length (x-axis) varies between 500 and 5,000 with a step of 500. Red line represents the true value of $H_{xy} = 0.7$.

For the power law coherency case, we use the mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ with perfectly correlated inno-

vations ε_2 and ε_3 and with uncorrelated other pairs to arrive at the expected bivariate Hurst exponent of 0.7. Figure 5.2 shows the results of simulations for the same parameter setting of the estimators as for the previous case. Bias-wise, the results are very straightforward – DMCA dominates the other two estimators while for DMCA and HXA, the mean value is practically unaffected by the time series length. For DCCA, the bias even increases with T which is very undesirable. DMCA also dominates the other two estimators in variance and thus also in the mean squared error comparison. The dependence of variance of the estimators with respect to the time series length T again approximately follows a power law.

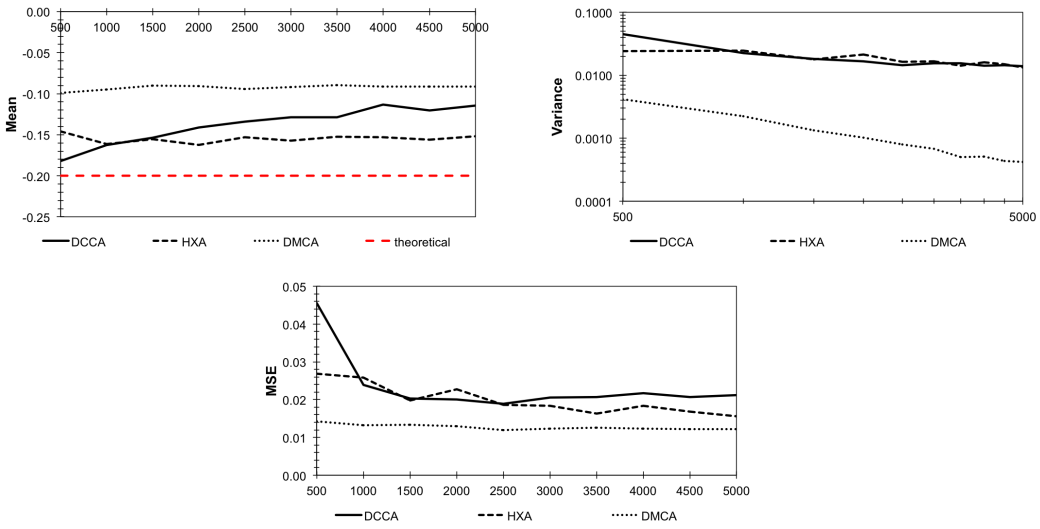


Figure 5.3: **Comparison of DCCA, HXA and DMCA power law coherency estimators.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$, $d_2 = d_3$ and perfectly correlated innovations ε_2 and ε_3. Time series length (x-axis) varies between 500 and 5,000 with a step of 500. Red line represents the true value of $H_\rho = -0.2$.

Apart from estimation of H_{xy} for the power law coherency case, we have redefined the time domain estimators for estimation of parameter H_ρ. Using the same specification as for the previous case, we present Figure 5.3. DMCA shows stable upward bias of approximately 0.1 regardless the time series length. The bias even increases with the time series length for DCCA. But for HXA, we see a very slow convergence to the true value of -0.2, specially for $T = 5000$, the estimator stabilizes at an upward bias of 0.05. With respect to variance, DMCA clearly outperforms the other two estimators. HXA and DCCA estimators attain very similar levels of variance as well as a decay. Putting bias

and variance together, the mean squared error shows a dominance of DMCA. However, the mean squared error of HXA gets very close to the DMCA levels and for longer time series, HXA might outperform DMCA even in MSE. The levels for DCCA stabilize for higher time series lengths and remain well above the other two estimators.

Based on the comparison presented in this section and on the finite sample properties studied in the previous section, it is evident that each of the three time domain estimators is better suited for different specifications of processes. DCCA and HXA are able to deal with short-term memory bias quite well and even for very strong short-term memory, these two estimators remain rather weakly biased. DMCA estimator of H_{xy} dominates the other two for the power law coherency case both in bias and variance. However, for the estimation of the power law coherency parameter H_ρ, HXA dominates in bias and gives mean squared errors comparable with DMCA, which shows remarkably low values of variance. In the next section, we present three frequency domain estimators and in the same steps as for the time domain estimators, we discuss their finite sample properties and compare their performance.

5.2 Frequency domain estimators

Frequency domain estimators are based on Proposition 2.6 which states that at frequencies close to the origin, the magnitude of the cross-power spectrum follows a power law and diverges to infinity. Estimation of the cross-power spectrum thus becomes crucial. The most frequently used tool is a cross-periodogram $I_{xy}(\lambda)$ defined as

$$I_{xy}(\lambda_j) = \frac{1}{2\pi T} \sum_{k=1}^{T} x_t \exp(-i\lambda_j k) \sum_{l=1}^{T} y_t \exp(i\lambda_j l) = I_x(\lambda_j)\overline{I_y(\lambda_j)}, \qquad (5.14)$$

where T is the time series length and λ_j is a frequency defined as $\lambda_j = 2\pi j/T$ where $j = 1, 2, \ldots, \lfloor T/2 \rfloor$, where $\lfloor\rfloor$ is the nearest lower integer operator so that the cross-periodogram is defined between 0 and π. $I_x(\lambda_j)$ is a periodogram of series $\{x_t\}$ and $\overline{I_y(\lambda_j)}$ is a complex conjugate of a periodogram of series $\{y_t\}$. If the cross-periodogram is used in its raw form, then evidently, we will always obtain $H_{xy} = \frac{H_x+H_y}{2}$ (based on Proposition 2.6). Moreover, the raw cross-periodogram (as well as the raw univariate periodogram) is an inconsistent estimator of the true cross-power spectrum (Wei 2006). To overcome

the inconsistency issue, the raw (cross-)periodogram needs to be smoothed. Bloomfield (2000) suggests a simple smoothing operator proposed by Daniell (1946), which is practically a simple moving average with half weights on the boundary values. Some authors (Velasco 1999; Hurvich & Chen 2000; Sela & Hurvich 2012) also suggest to first taper the series to deal with leakages at low frequencies. In the following text, we apply only the smoothing of periodograms while the tapering is not utilized as it did not show any finite sample efficiency or bias gains for the estimators we use.

5.2.1 Averaged periodogram estimator

Sela & Hurvich (2012) propose the averaged periodogram estimator (APE) and they are in fact the first ones to propose an estimator of H_{xy} (or more correctly d_{12} in their case) in the frequency domain. The estimator is a bivariate generalization of the method of Robinson (1994). Taking the cumulative cross-periodogram $\widehat{F_{xy}}(\lambda) = \frac{2\pi}{m}\sum_{j=1}^{\lfloor m\lambda/2\pi \rfloor} I_{xy}(\lambda_j)$ where $m \leq T/2$ is a bandwidth parameter and fixed $q \in (0,1)$, the estimator is given by

$$\widehat{H_{xy}} = 1 - \frac{\log \frac{\widehat{F_{xy}}(q\lambda_m)}{\widehat{F_{xy}}(\lambda_m)}}{2\log q}. \tag{5.15}$$

Under twelve assumptions given by Sela & Hurvich (2012), the estimator is consistent. Moreover, it is advised to use $q = 0.5$. The authors also provide a Monte Carlo simulation study to show the finite sample properties of the estimator. The bias and efficiency are shown by box plots for several scenarios to show that for samples below 10,000 observations, the estimator is strongly biased with high variance. For high values of m and high number of observations, the variance of the estimator decreases markedly while the estimator still remains biased. To study the properties of the estimator in more detail, we provide a wide Monte Carlo study at the end of this chapter where APE is compared with the other two estimators.

5.2.2 Cross-periodogram estimator

As proved in Proposition 2.6, the cross-power spectrum diverges at the origin as a power law with exponent β_{xy}. As the previous estimators focus on the estimation of the bivariate Hurst exponent H_{xy}, we can rewrite the scaling of

the cross-power spectrum using Proposition 2.7 as

$$|f_{xy}(\lambda_j)| \propto \lambda_j^{-(2H_{xy}-1)}. \tag{5.16}$$

Using the cross-periodogram as an estimator of the cross-power spectrum, we expect the long-range cross-correlated series to follow

$$|I_{xy}(\lambda_j)| \propto \lambda_j^{-(2H_{xy}-1)}. \tag{5.17}$$

The cross-periodogram estimator (XPE) of the bivariate Hurst exponent H_{xy} can be obtained through a regression on

$$\log |I_{xy}(\lambda_j)| \propto -(2H_{xy}-1)\log \lambda_j. \tag{5.18}$$

As the power-law scaling is expected only for $\lambda \to 0+$, the regression is not performed over all frequencies. By choosing $\lambda_j = 2\pi j/T$ for $j = 1, 2, \ldots, m$ where $m \leq T/2$, we estimate the bivariate Hurst exponent using only the information up to a selected frequency. For the univariate case, Beran (1994) and Robinson (1995b) show that the periodogram estimator is consistent and asymptotically normal with

$$\sqrt{m}(\widehat{H} - H^0) \to_d N(0, \pi^2/24) \tag{5.19}$$

where H^0 is the true Hurst exponent. The limiting distribution is free of H^0 and also of all the other parameters (the assumptions are given in Beran (1994) [Theorem 4.6]). The variance of the estimator decreases with a square root of m but we need to keep in mind that the higher the m parameter is, the more biased the estimator is because the power-law scaling holds only for the origin neighborhood. Choice of m thus depends on preferences between bias and efficiency. For the bivariate case, however, we have more parameters in the specification of the model – mainly the univariate Hurst exponents of the separate processes and the correlation coefficient between innovations for the simplest case – and there is no reason to believe that the properties of the XPE estimator would be independent of these. Showing the asymptotic properties requires a strict set of assumptions and the underlying bivariate model specification. In the context of this thesis, it would be out of line to assume some particular specification of the underlying model as we try to keep the assumptions of the methods as general as possible. We thus do not provide

the asymptotic properties for this estimator and leave it for further research, yet still, we provide a discussion about dependence of mean and variance of the estimator on the set of parameters and compare the properties with the other two frequency-based estimators in Subsection 5.2.6.

5.2.3 Local X-Whittle estimator

Local Whittle estimator of the fractional differencing parameter d or Hurst exponent H is based on the same principle as the previously defined periodogram estimator – the power-law divergence of the power spectrum. However, instead of the regression fitting to the power-law scaling near the origin as $\lambda \to 0+$, local Whittle estimator is based on minimization of the penalty function based on Künsch (1987).

Taking Robinson (1995a) as a starting point and generalizing the method for the bivariate series, we propose the estimator of the bivariate Hurst exponent H_{xy} as follows. Divergence of the magnitude of the cross-power spectrum close to the origin with the power-law scaling according to Proposition 2.6 is assumed for long-range cross-correlated processes $\{x_t\}$ and $\{y_t\}$. Cross-periodogram $I_{xy}(\lambda)$ is defined according to Equation 5.14 with $j = 1, 2, \ldots, m$ where $m \leq T/2$ and $\lambda_j = 2\pi j/T$. Assuming that series $\{x_t\}$ and $\{y_t\}$ are indeed long-range cross-correlated with $\frac{1}{2} < H_{xy} \leq 1$, we propose the local X-Whittle estimator (LXW) as

$$\widehat{H_{xy}} = \arg \min_{\frac{1}{2} < H_{xy} \leq 1} R(H_{xy}), \tag{5.20}$$

where

$$R(H_{xy}) = \log \left(\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2H_{xy}-1} |I_{xy}(\lambda_j)| \right) - \frac{2H_{xy}-1}{m} \sum_{j=1}^{m} \log \lambda_j. \tag{5.21}$$

and $\lambda_j = 2\pi j/T$. The LXW estimator is thus a semi-parametric maximum likelihood estimator as it utilizes only the properties of the cross-power spectrum near the origin.

In a similar manner as for the XPE estimator, the univariate version of the local Whittle estimator is consistent and asymptotically normal, specifically

$$\sqrt{m}(\widehat{H} - H^0) \to_d N(0, 1/4) \tag{5.22}$$

where again H^0 is the true bivariate Hurst exponent and the limiting distribu-

tion is free of H^0 and all the other parameters. The local Whittle estimator is thus asymptotically more efficient than periodogram estimator as $1/4 < \pi^2/24$. For detailed treatment of the univariate case and the assumptions, see Robinson (1995a). In the bivariate case, there is again no reason to presume that the asymptotic properties would be independent of the univariate Hurst exponents and the correlation structure of the innovations. Discussion of these possible dependencies and comparison with APE and XPE is given in Subsection 5.2.6.

5.2.4 Power law coherency estimation

Apart from estimating the bivariate Hurst exponent H_{xy}, the frequency domain is appropriate for estimating the power law coherency and more precisely, estimating whether the two processes are power law coherent. As the frequency domain is closer to the actual definition of the power law coherency, we expect that the frequency-based estimators of such coherency would be more accurate than the time domain ones. This is discussed later in the section.

Starting from the ideas of Sela & Hurvich (2012), we propose a set of estimators which assume that the estimators of the univariate and bivariate Hurst exponents are consistent. Sela & Hurvich (2012) show that this is a case for APE under rather strict assumptions (see the reference for the list). In Subsection 5.2.5-5.2.6, we argue that also XPE and LXW are consistent estimators of H_{xy} even though it is based only on the results of Monte Carlo simulations. The univariate versions of XPE and LXW are consistent for H_x and H_y as argued in Beran (1994) and Robinson (1995a;b). It thus seems convenient to propose a simple estimator of the power law coherency H_ρ as

$$\widehat{H_\rho} = \widehat{H_{xy}} - \frac{\widehat{H_x} + \widehat{H_y}}{2}. \tag{5.23}$$

We thus get the estimator of H_ρ based on the separate estimators of H_{xy}, H_x and H_y for APE, XPE and LXW labeling them as $\widehat{H_{\rho,APE}}$, $\widehat{H_{\rho,XPE}}$ and $\widehat{H_{\rho,LXW}}$, respectively.

Apart from the estimators based on Equation 5.23, we also propose estimators in the steps of XPE and LXW. As both estimators are based on the power-law scaling of the cross-power spectrum, parallel estimator can be constructed for the power-law scaling of the squared spectrum coherence. Note that the scaling of the squared spectrum coherency is given (see proof to Proposition 2.8)

as

$$K_{xy}^2(\lambda) = \frac{|f_{xy}(\lambda)|^2}{f_{xx}(\lambda)f_{yy}(\lambda)} \propto \frac{\lambda^{2(1-2H_{xy})}}{\lambda^{1-2H_x}\lambda^{1-2H_y}} = \lambda^{-4(H_{xy}-\frac{H_x+H_y}{2})} = \lambda^{-4H_\rho}. \quad (5.24)$$

Using the cross-periodogram as an estimator of the cross-power spectrum, we can reformulate the XPE for the power law coherency estimate $\widehat{H_{\rho,XPE}}^*$ in a sense of the log-log regression on

$$\frac{|I_{xy}(\lambda_j)|^2}{I_x(\lambda_j)I_y(\lambda_j)} \propto \lambda_j^{-4H_\rho} \quad (5.25)$$

where again $\lambda_j = 2\pi j/T$ and $j = 1, \ldots, m$ with $m \leq T/2$.

In a similar manner, we can reformulate the LXW estimator so that the power law coherency can be estimated as

$$\widehat{H_{\rho,LXW}}^* = \arg\min_{H_\rho \leq 0} R(H_\rho), \quad (5.26)$$

where

$$R(H_\rho) = \log\left(\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{4H_\rho} \frac{|I_{xy}(\lambda_j)|^2}{I_x(\lambda_j)I_y(\lambda_j)}\right) - \frac{4H_\rho}{m}\sum_{j=1}^{m}\log\lambda_j. \quad (5.27)$$

and $\lambda_j = 2\pi j/T$, $j = 1, \ldots, m$ and $m \leq T/2$.

In the following two sections, we provide an extensive discussion of the finite sample properties of the frequency-based estimators.

5.2.5 Finite sample properties

Compared to the time domain estimators, where we have been interested mainly in the effect of various scales parameters (s_{min}, τ_{max} and κ_{max}) on their statistical properties, we are interested in the effect of smoothing cross-periodogram and a part of cross-periodogram (bandwidth parameter m) used for the estimation. For smoothing, we compare the values of Daniell moving averages of lengths 11, 21 and 51. We do not deal with a raw cross-periodogram, which would push the estimates of the bivariate Hurst exponent to the average of the separate Hurst exponents. As for the bandwidth parameter m, we compare two cases – $m = 0.1T$ and $m = 0.4T$ – which translates into the maximum frequencies used of $\lambda_{max} = 0.2\pi$ and $\lambda_{max} = 0.8\pi$, respectively. We thus compare the bias and efficiency gains (and losses) based on varying the parameters.

Starting with the APE method, we observe that the choice of the above mentioned parameters influences the statistical properties significantly. For the correlated ARFIMA processes (Table C.22-C.23), we find that bias-wise, APE is quite similar to the time domain estimators – for weak long memory, the estimator is only slightly biased but for strong long memory, the estimator is downward biased. The choice of m parameter has an interesting effect – $m = 0.4T$ provides lower bias and lower variance for all combinations of other parameters. The effect of smoothing varies across time series lengths – for shorter series, the higher level of smoothing brings higher bias and lower variance which, however, results in the increasing mean squared error with the smoothing length. For longer series, $T = 5000$, the bias and variance do differ much for different levels of smoothing in most cases, even though there are differences for the case of low correlation between innovations.

For ARFIMA and AR(1) based long-range cross-correlated processes (Table C.24-C.26), we observe an expected difference for $m = 0.1T$ and $m = 0.4T$. For the case of a weak short memory (Table C.24), the statistical properties do not differ much compared to the case of correlated ARFIMA processes with $d = 0.4$ discussed in the previous paragraph. However, when the strength of short memory increases, the situation changes. For $\theta = 0.5$, the $m = 0.1T$ setting outperforms the $m = 0.4T$ as the bias caused by short memory is not present in the very low frequencies of the periodogram but already affects the frequencies taken into consideration in the latter case. Even though the variance is obviously lower for the latter case, the total effect captured by the mean squared error clearly shows that the former parameter choice is superior. The choice of smoothing parameter affects the statistical properties as well – the estimates are pushed lower with the increasing level of smoothing and the variance decreases in most cases, even though the effect on variance is very mild for processes with highly correlated innovations. For strong short memory ($\theta = 0.8$), the effect on periodogram is so strong that even the $m = 0.1T$ based estimates are strongly biased upwards. Other observations remain the same as for the previous case. Comparing the total effect of the short memory bias for this last case, the $m = 0.1T$ setting brings approximately three times lower mean squared error compared to the $m = 0.4T$ setting. The efficiency gain of the latter case is thus more than offset by the increase of bias.

For the case of the mixed-correlated ARFIMA processes (Table C.27), we again find practically the same behavior as for the previous cases – smoothing increases the variance and pushes the estimates lower while the bandwidth

parameter decreases the variance of the estimator as expected. The effect of smoothing on bias seems to be dependent on the time series choice. Further research on this matter should be made (and is partially discussed in Sela & Hurvich (2012)) but is out of scope of this thesis. Nonetheless, for a correct choice of parameters, the estimator is unbiased for the processes with high level of correlations, i.e. highly correlated innovations. The bias is thus also dependent on the correlation between innovations – it decreases with the level of correlation.

As an estimator of the power law coherency, the APE estimator brings ambiguous results (Table C.28). First, the estimator is biased upwards for all cases. Even though for the low levels of correlation between innovations, the estimator outperforms the time domain estimators and the opposite is true for the higher levels of correlations. Second, the choice of the bandwidth parameter m brings both bias and efficiency gains. Third, stronger smoothing of periodogram brings lower bias but higher variance to the estimators. Fourth, the estimator does not show any evident convergence to the true value of the parameter as the bias remains practically unchanged for various time series lengths. Keeping in mind that it has been shown by Sela & Hurvich (2012) that the estimator is consistent, the rate of convergence seems to be very slow. Even for the longest analyzed series of 5,000 observations, the best case turns out to be a bias of 0.11, which is not very desirable. However, compared to the time domain estimators, the variance remains on lower levels.

Finite sample properties of XPE are actually similar to the ones of APE and are summarized in Table C.29-C.35. For correlated ARFIMA processes (Table C.29-C.30), the estimators can be seen as only slightly downward biased apart from the shortest analyzed series for which the bias is more severe. Nonetheless, the bias and variance decreases with the time series length which hints a consistence of the estimator. In most of the cases, increasing the smoothing level increases the variance of the estimator. Compared to APE, XPE yields less biased and more efficient estimates of the bivariate Hurst exponent for the correlated ARFIMA processes, which is reflected in the mean squared errors which are approximately 2.5 times higher for APE compared to XPE.

The short-range dependence bias is illustrated in Table C.31-C.33 and the results again follow the same logic as for APE. For the weak short memory, $\theta = 0.1$, there is practically no difference compared to the strongly dependent correlated ARFIMA processes in the previous paragraph. Increasing the strength of the short memory is reflected in the bias of the estimators. For a

medium memory, the bias is evident for $m = 0.4T$ while for $m = 0.1T$, the bias increases only mildly. For the strong short-term memory, both specifications lead to strongly biased estimates. Even though the variance is lower for the $m = 0.4T$ case, the total effect of the short memory leads to higher mean squared errors compared to the case of $m = 0.1T$. Increasing the Daniell windows leads to less biased estimates but higher variance. The total effect on the mean squared error varies across correlations between innovations and among the time series lengths. Compared to the APE, the mean squared errors are approximately twice higher for XPE for strongly short-term memory influenced series. APE thus seems to be more prone to the short-term memory bias.

Properties of the estimator for the mixed-correlated ARFIMA processes are summarized in Table C.34. We again observe similar qualitative properties as for APE. First, higher level of smoothing pushes the estimates downwards and increases the variance of the estimator. Second, the increasing correlation between innovations pushes the estimates downwards but has no monotone effect on variance. Third, increasing m lowers both variance and bias in most cases. Probably most importantly, the bias does not decrease with time series length so that based on the simulations, the estimator does not seem to be consistent. However, such statement holds only for the case when the smoothing parameter is fixed and not related to the time series length in any way, which might not be an optimal choice. Compared to APE, the mean squared errors are similar for the XPE method.

The power law coherency is estimated by two estimators based on XPE, according to Equation 5.23 and Equation 5.25. The results are summarized in Table C.35 and Table C.36, respectively. Both estimators share some properties. First, smoothing decreases the bias but increases the variance in most cases. Second, higher m yields lower variance but also higher bias in majority of cases. Third, the bias does not decrease with the time series length monotonously. The estimators are thus either inconsistent or the rate of convergence is very slow. Nonetheless, the estimator based on the scaling of squared spectrum coherence performs better than the one based on the difference between the bivariate Hurst exponent and the average of the separate Hurst exponents – the coherence-based estimator is less biased and the mean squared errors are lower than the respective ones for the latter estimator.

LXW estimator follows the similar pattern as APE and XPE for the correlated ARFIMA processes (Table C.37-C.38) – practically unbiased for the weak long memory and only slight downward bias for the strong long memory. The

effect of smoothing is not monotonous for the estimator but for the majority of cases, the variance increases with the smoothing length. Bandwidth parameter m decreases the variance of the estimator and the effect on bias is not evident in the analyzed finite samples.

The effect of short-term memory is again similar to the previous two estimators as shown in Table C.39-C.41. For the weak short-term memory, the properties of the estimator remain practically the same as for the case of strong long-term memory. The medium short-term memory has an effect on the estimates with $m = 0.4T$ while the ones of $m = 0.1T$ remain quite stable. The mean squared errors are comparable to the ones of APE and XPE. For the strong short-term memory, the LXW estimates become severely upward biased for both levels of m analyzed here. Compared to the other two frequency domain estimators, LXW seems to be the one most biased by the presence of the short-term memory component in the processes.

The results for the simulations based on the mixed-correlated ARFIMA processes are summarized in Table C.42. We observe that the variance of estimator increases with the smoothing strength and decreases with the strength of correlation between innovations. For the bias, the relationship is not so clear. Higher m brings lower variance of the estimator but does not necessarily lead to lower bias. Apart from the longest analyzed series, the mean squared error majorly increases with the smoothing parameter. The total performance of the estimator is comparable to the other two frequency domain methods based on the mean squared error.

For the estimation of the power law coherency parameter, we again use two estimators as for the XPE method (Table C.43-C.44). There are again several regularities which are common for both estimators. First, they are strongly upward biased. Second, the bias decreases with the smoothing parameter. Third, the variance increases with the smoothing parameter. Fourth, the bias decreases with the growing correlation of the innovations. The same is true for the variance in most cases. Comparing the two methods of estimation, the coherence-based estimator is more biased but has lower variance compared to the estimator based on the difference between estimated bivariate Hurst exponent and the average of the separate Hurst exponents. Based on the mean squared error, the latter estimator is preferred for majority of the cases. However, the differences in the mean squared errors are usually minor.

5.2.6 Dependence on bandwidth parameter

Choice of the bandwidth parameter m is a crucial aspect of the frequency domain estimators as shown in the previous section. As noted in the studies dealing with the univariate specifications of the estimators (Beran 1994; Robinson 1995a;b), variance should decrease with the parameter and bias should increase. The former comes from the fact that the estimation is based on more data points and latter from the fact that the power-law scaling of the cross-power spectrum holds only for the lowest frequencies. In this section, we present and discuss the behavior of the mean and variance of the frequency-based estimators presented above with respect to varying parameter m.

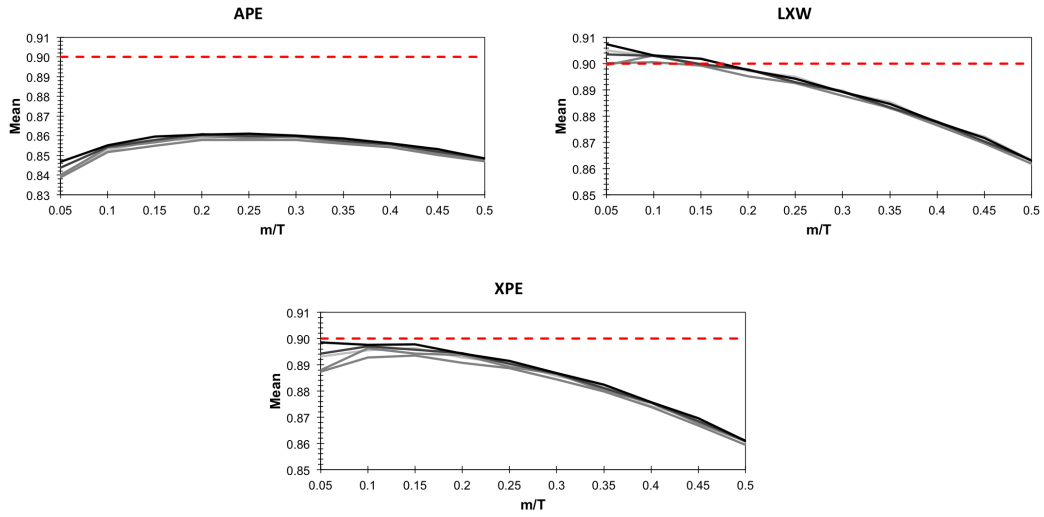


Figure 5.4: **Mean values of APE, LXW and XPE estimators dependent on m and correlation between innovations I.** Values are based on 1,000 simulations of ARFIMA(0,d,0) processes with correlated innovations and $d_1 = d_2 = 0.4$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1). Red line represents the true value of $H_{xy} = 0.9$.

We discuss two main scenarios of the long-range cross-correlated processes – the processes with and without power law coherency behavior. For the latter, we utilize ARFIMA(0,d,0) processes with correlated innovations and with $d_1 = d_2 = 0.4$ and for the former, we use the mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$. Both kinds of processes are studied for the longest time series length of the previous section – $T = 5000$ – and correlation between innovations varying between 0.2 and 1 with a step of 0.2. To uncover the dependence on m, we use m/T from 0.05 up to 0.5 with a

step of 0.05. We thus cover the cross-periodogram from the lowest tenth of the frequencies up to the whole cross-periodogram. For each specification, we use 1,000 simulations and the Daniell's window of 21 as used in Sela & Hurvich (2012). This way, we are able to comment on the dependence of bias and variance of the estimator with respect to the correlation between innovations and the bandwidth parameter m.

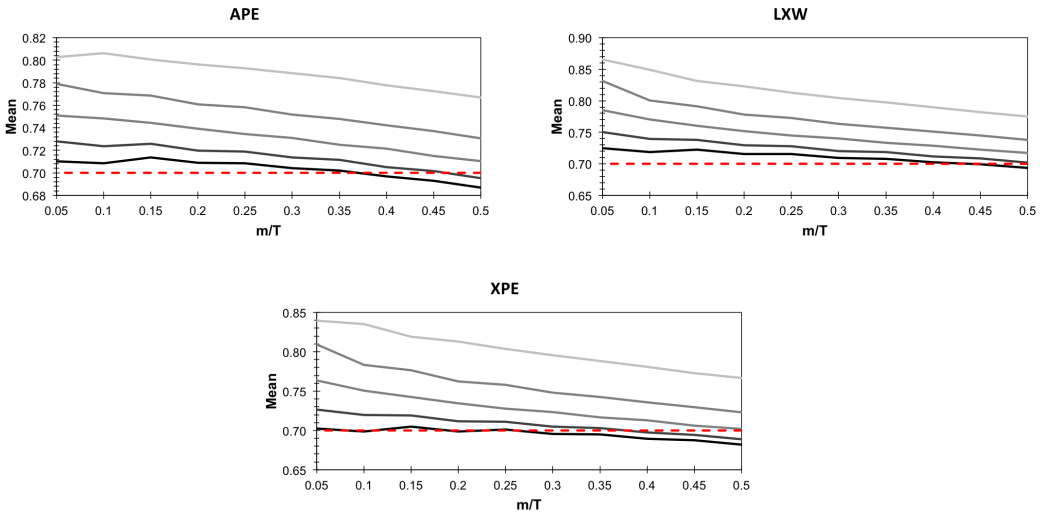


Figure 5.5: **Mean values of APE, LXW and XPE estimators dependent on m and correlation between innovations II.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1). Red line represents the true value of $H_{xy} = 0.7$.

Starting with the bias, Figure 5.4 and Figure 5.5 show the results of simulations for the correlated and mixed-correlated ARFIMA processes. In Figure 5.4, we observe that for the correlated ARFIMA processes for which we expect the bivariate Hurst exponent to be equal to the average of the separate Hurst exponents, the bias behavior differs for specific methods. For LXW and XPE, we observe an expected behavior – the estimates are unbiased for approximately $m/T \leq 0.2$. For higher values of m, the estimates become biased downwards. Interestingly, the bias is practically independent of the correlation level between innovations for all three estimators. However, for APE, the mean values of the estimates are very stable across various m but remain well below the theoretical value of 0.9 and yield a negative bias of approximately -0.05. Again, the bias is practically independent of the correlations between innovations.

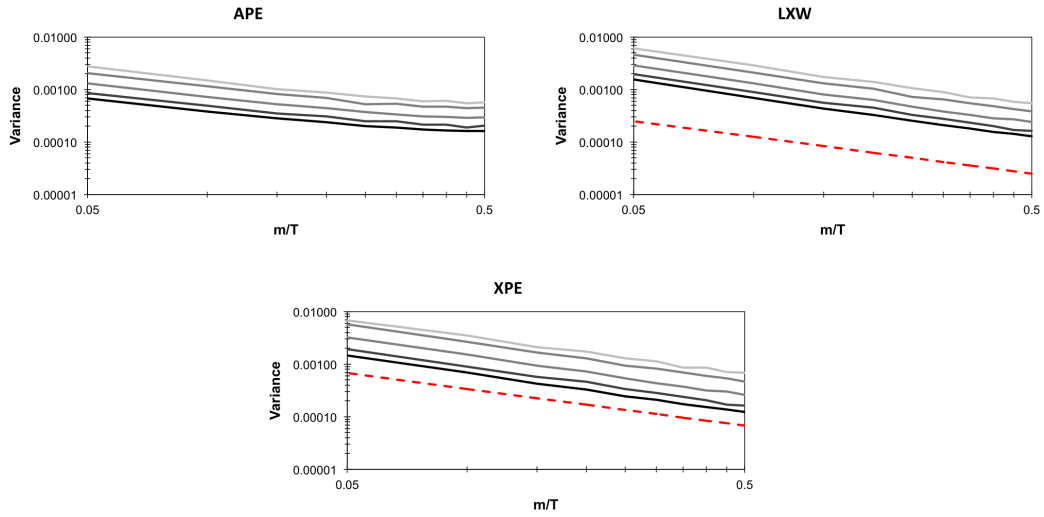


Figure 5.6: **Variance of APE, LXW and XPE estimators dependent on m and correlation between innovations I.** Values are based on 1,000 simulations of ARFIMA(0,d,0) processes with correlated innovations and $d_1 = d_2 = 0.4$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1). Red line represents the asymptotical values for the univariate case (for LXW and XPE only).

The situation is more interesting for the mixed-correlated ARFIMA processes, i.e. the power law coherency case. In Figure 5.5, we can see that the mean values of the estimates are dependent on both m and the correlation between innovations for all three estimators. In general, it holds that the estimates are less biased with an increasing correlation between innovations, which is expected, but also with higher m, which is well in hand with the finite sample properties of the estimators presented in the previous section. The performance of the estimators is thus not only dependent on the parameters but also on the specification of the model as shown by the difference between the two cases.

The situation is quite similar for the behavior of variance of the estimators. In Figure 5.6, we show this behavior compared to the theoretical asymptotic variance for the univariate cases (for LXW and XPE). We observe several regularities which are true for all three estimators. First, the variance decreases with increasing m as expected. Second, the variance decreases with increasing strength of correlations between innovations. Third, the log-log depiction indicates a power-law scaling with parameter m. For LXW and XPE, we can compare this scaling with the asymptotic scaling for the respective univariate estimators and it is visible that these can be seen as square-root scalings which

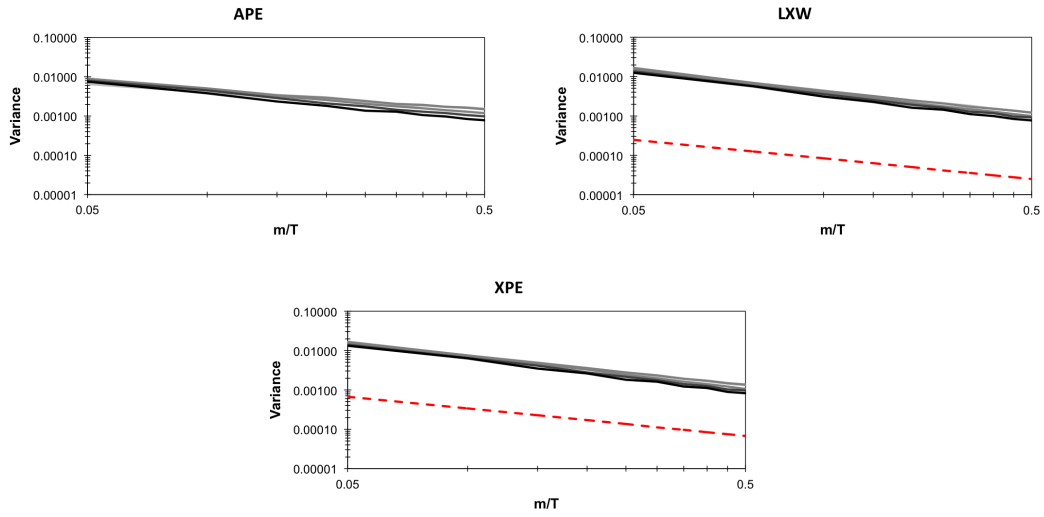


Figure 5.7: **Variance of APE, LXW and XPE estimators dependent on m and correlation between innovations II.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1).

are in hand with the univariate case. However, the levels of variance are well above the asymptotic univariate values, which is, however, not surprising. The variance levels of LXW and XPE practically overlap while the variance of APE shows slightly lower values.

For the mixed-correlated ARFIMA processes, Figure 5.7 depicts the behavior of variance. All the regularities from the previous paragraph hold even here. However, the level of variances is evidently higher for this case and shows that the estimators have much higher variance for the power law coherency specification. The distance from the univariate asymptotic variances of LXW and XPE is much more profound as well. Again, the level of variances is lower for APE compared to the other two estimators.

Focusing now on the estimators of the power law coherency, we use the same specifications of the processes as well as of the estimators. In Figure 5.8, we show mean values of the estimated power law coherency for both specifications of the LXW and XPE estimators and for the APE estimator. We observe that the mean values are very stable across various m and that the bias decreases with increasing correlation between innovations in the mixed-correlated ARFIMA specification. Nevertheless, the bias remains quite severe even for perfectly correlated innovations. The best performance in the bias

sense is found for XPE coherence-based estimator which gets to a bias of 0.1. Evidently, the estimators would need much longer series to perform adequately which is in hand with results presented in Sela & Hurvich (2012) for APE.

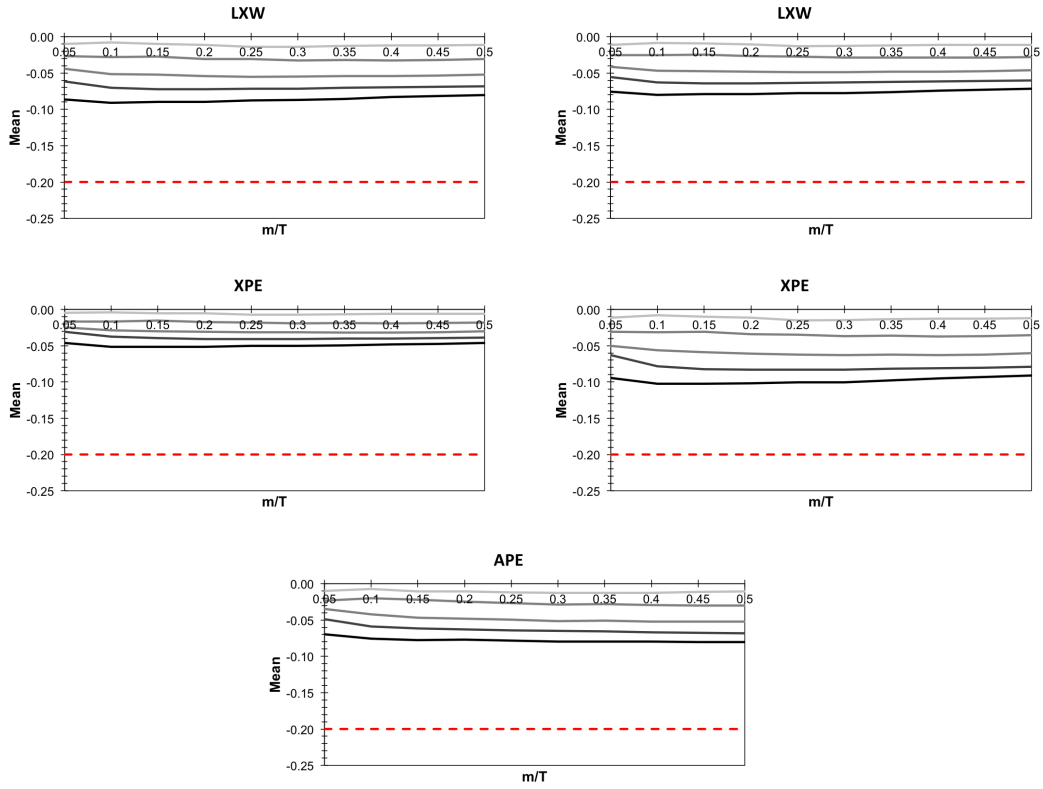


Figure 5.8: **Mean of APE, LXW and XPE estimators of power law coherency dependent on m and correlation between innovations.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1). Red line represents the true value of $H_\rho = -0.2$. For LXW and XPE which have two possible definitions of the power law coherency, the coherency-based estimators are shown on the right and the estimators based on the bivariate and separate Hurst exponents are shown on the left.

In Figure 5.9, we show the dependence of variance of the estimators on m. Interestingly, the variance is practically independent of the correlation between innovations. Expectedly, the variance decreases with increasing m and specifically in a power-law manner. For LXW, the values practically overlay for both specifications. For XPE, the variance of the estimator based on the bivariate and separate Hurst exponents shows lower levels of variance than the coherence-based one. APE estimator attains levels of variance similar to the

others except for the Hurst exponent based XPE which reaches markedly lower values.

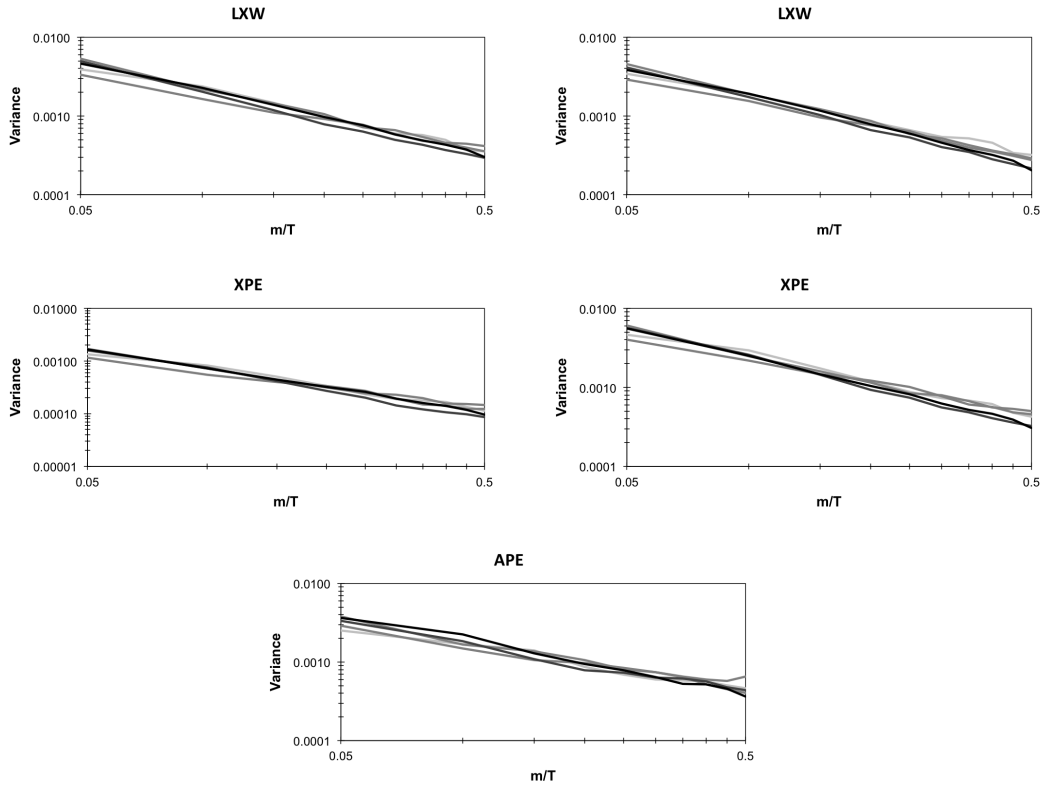


Figure 5.9: **Variance of APE, LXW and XPE estimators of power law coherency dependent on m and correlation between innovations.** Values are based on 1,000 simulations of mixed-correlated ARFIMA(0,d,0) processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$. Correlation between innovations ranges between 0.2 and 1 with a step of 0.2 and is represented by different shades of grey in the chart (the lightest for 0.2 and black for 1). For LXW and XPE which have two possible definitions of the power law coherency, the coherency-based estimators are shown on the right and the estimators based on the bivariate and separate Hurst exponents are shown on the left.

5.3 Brief overview

In this chapter, we have introduced six estimators of the bivariate Hurst exponent and we have discussed their finite sample properties as well as their other characteristics. The results are quite expected and the most important ones are summarized below.

First, the frequency domain estimators outperform the time domain esti-

mators in estimating H_{xy} in both bias and variance if the specified process comprises of two long-range dependent processes, i.e. the properties of the cross-power spectrum at origin are not influenced by short-term memory at higher frequencies. However, if the short-term memory is part of the underlying bivariate process, the frequency-based estimators become strongly biased and practically break down for strong short-term memory. The trade-off between bias and variance for various values of m plays a crucial role here.

Second, the time domain estimators are less efficient than the frequency domain ones in most cases which is, however, not surprising. On contrary, DCCA and HXA are quite robust to short-term memory bias. Probably the most severe issue of the time domain estimators is the fact that between their two specifications – with and without the absolute values – there is a big difference. For the specifications without the absolute values, the estimators are able to actually estimate anything only for rather high levels of correlation between processes because for the lower values, it frequently happens that negative covariances become a part of the final regression which breaks down the power-law scaling assumptions. Reversely, the absolute value based estimators have no problem with power-law scaling assumption but the absolute value specification pushes the estimates of the bivariate Hurst exponent to the average of the separate Hurst exponents regardless the actual value of H_{xy}. This makes time domain estimators quite hard to use for the bivariate Hurst exponent estimation if the bivariate exponent is lower than the average of the separate ones. Unfortunately, the power law coherency is potentially the most interesting case of the long-range cross-correlations.

Third, and quite ironically, time domain estimators of the power law coherency parameter H_ρ, and namely HXA, are much less biased than the ones in frequency domain. This leads to an interesting situation when the time domain estimators should be used to uncover whether the power law coherency is present between the analyzed processes but for the actual estimation of H_{xy}, we should use the frequency domain estimators. Nonetheless, this creates rather interesting situation and we cannot simply discard one of the approaches but these should be utilized together.

Chapter 6

Leverage effect between financial returns and volatility

A negative relationship between returns and changes in volatility is a well-documented phenomenon in the financial economics. Already Black (1976) discussed a possible relationship between returns and changes in volatility. The argumentation was made on a basis of changes in earnings. If expected earnings of a company go down, the price goes down as well creating negative returns. Decreased expected earnings draw a value of the company down and in turn increase the leverage (ratio between debt and equity) resulting in higher riskiness of holding the stock which comes with higher volatility as a synonym to risk. The effect is thus usually referred to as the 'leverage effect'. Black also briefly discussed the possibility of an inverse causality coming from a stable business (low volatility) to better future prospects (positive returns) and vice versa. However, he argues that only the former causality (from returns to volatility) makes sense in the efficient markets hypothesis framework (Fama 1965; 1970; Samuelson 1965) because if the latter held, it would be used for obtaining abnormal profits as the volatility information is publicly available. The former causality might, however, hold true as predicting volatility does not bring abnormal returns. Similar logics is investigated by various authors such as Bekaert & Wu (2000), Carr & Wu (2009), Ismail (2011) and Smith & Yamagata (2011).

Investigation of the leverage effect is frequently connected to the asymmetric volatility phenomenon, which is usually treated as a situation when volatility of a growing market tends to be lower than volatility of a falling market. The interconnection is thus very tight and it is mostly quite hard to distinguish

between these two effects. Nonetheless, most of the authors agree on several regularities – correlation between returns and volatility is negative but rather weak and the effect comes from the returns to volatility and lasts for several periods while remaining negative and quite persistent (Pagan 1996; Bouchaud & Potters 2001; Bouchaud *et al.* 2001; Bollerslev *et al.* 2006). These regularities led to a proposition of several models which can capture and mimic them. These are mainly the diffusion-based processes (Bouchaud & Potters 2001; Bouchaud *et al.* 2001; Bollerslev *et al.* 2009; Li 2011) and the generalizations of (G)ARCH family processes (Pagan & Schwert 1990; Bekaert & Wu 2000; Martens *et al.* 2009; Wang & Yang 2009; Hammoudeh *et al.* 2010; Talpsepp & Rieger 2010; Ismail 2011). Recently, Bollerslev *et al.* (2012) propose an internally consistent equilibrium model which explains the leverage effect as well as several other stylized facts of the financial markets. Aït-Sahalia *et al.* (2013) provide a nice review of literature on the topic and focus on a specific form of the leverage effect – relationship between changes in volatility and returns – finding a puzzle of a weak evidence of this kind of the leverage effect. They show that this is caused by a high-frequency bias of the volatility estimators (mainly microstructure noise and jumps) which are used in the analysis. In effect, the leverage effect might not be found purely due to this bias. After controlling for this bias, the authors show that the negative correlation between returns and changes in volatility are observed. We try to tackle the phenomenon from a different viewpoint utilizing the methodology proposed in the text up till this point.

Main aim of this chapter is to approach the leverage effect from the long-range cross-correlations perspective and to possibly find some implications which can be derived from the empirical data. We do not intend to propose a new theoretical model but we are rather interested in a possibility that the leverage effect is a built-in feature of the data-generating process. To this point, we use two stylized facts about the financial returns – no serial correlation between returns and long-range dependence of volatility regardless the utilized volatility measure (Cont 2001; Poon & Granger 2003; 2005). The leverage effect thus becomes an ideal candidate for a long-range cross-correlations inspection.

We focus on 14 stock indices and analyze the leverage effect while focusing on the presence of long-range cross-correlations between returns and volatility. We then inspect whether the effect arises from the properties we find and whether it can be mimicked by a simple model that we propose. We first describe the dataset and its statistical properties. Then, we test whether the

processes of returns and volatility are long-range cross-correlated and since the majority of the analyzed indices turns out to be long-range cross-correlated, we follow with the analysis of the power law coherency. We find that returns and volatility do not resemble a power law coherent pair and the bivariate Hurst exponent thus does not differ from the average of the separate ones. This finding allows us to construct a relatively simple model that can mimic the leverage effect as well as several other stylized facts of the financial returns. In the framework of the proposed model, potentially spurious causality from returns to volatility is discussed in some detail.

6.1 Data description

We analyze the leverage effect between returns and volatility of 14 stock indices. The choice of indices is limited by data availability mainly due to the volatility series. As will be discussed later, we opt for the realized volatility as an estimate of volatility. However, the time series of the realized variance is publicly available only for a limited set of indices. We select 14 indices from the Oxford-Man Institute of Quantitative Finance website[1], which are summarized in Table 6.1. Indices from Europe (both the EU and non-EU), the Northern and Latin Americas and Asia are covered. The analyzed period ranges from 1.1.2000 to 31.12.2012 while the time series lengths obviously vary due to different national holidays.

Table 6.1: **List of analyzed stock indices.**

Label	Full name	Country	# of observations
AEX	Amsterdam Exchange Index	Netherlands	3306
BVSPA	Índice Bovespa	Brazil	3168
CAC	Cotation Assistée en Continu (CAC40)	France	3307
DAX	Deutscher Aktien IndeX (DAX30)	Germany	3290
DJIA	Dow Jones Industrial Average	USA	3242
EuroSTOXX	EURO STOXX 50 Index	EU	3282
FTSE	FTSE 100 Index	UK	3258
HSI	Hang Seng Index	Hong Kong	2938
IBEX	Índice Bursatil Español (IBEX35)	Spain	3272
KOSPI	Korea Composite Stock Price Index	South Korea	3201
NASDAQ	NASDAQ-100 Index	USA	3245
NIKKEI	NIKKEI 25 Index	Japan	3145
SPX	Standard & Poor's 500	USA	3240
SSMI	Swiss Market Index	Switzerland	3253

[1]http://realized.oxford-man.ox.ac.uk/data

For the returns part of the leverage effect analysis, we utilize logarithmic returns $r_t = \log(P_t) - \log(P_{t-1})$, where P_t is the price (value) of the analyzed index at time $t = 1, 2, \ldots, T$, and also the standardized returns $r_t^* = r_t/\widehat{\sigma}_t$ where $\widehat{\sigma}_t$ is an estimated standard deviation (volatility) at time t. We discuss the estimated volatility series later. The descriptive statistics of the returns series are summarized in Table 6.2. Apart from the standard statistics, which are in hand with the stylized facts – negative skewness for majority of indexes and excess kurtosis, i.e. fat tails, for all of them –, we also report two tests of normality – Jarque-Bera test (Jarque & Bera 1980; 1981) and Shapiro-Wilk test (Shapiro & Wilk 1965) – which show that normality is strongly rejected for all indices as expected (Cont 2001).

Table 6.2: **Descriptive statistics – returns.**

	mean	SD	skewness	excess kurtosis	Jarque-Bera	Shapiro-Wilk
AEX	-0.0005	0.0126	-0.1789	6.1016	5145.92***	0.9222***
BVSPA	0.0002	0.0186	-0.2118	4.7399	2989.23***	0.9620***
CAC	-0.0005	0.0130	-0.1197	3.9612	2169.95***	0.9489***
DAX	-0.0004	0.0142	-0.0706	4.5318	2818.04***	0.9470***
DJIA	0.0002	0.0124	0.0252	7.5007	7600.19***	0.9223***
EuroSTOXX	-0.0004	0.0143	-0.1202	4.1889	2407.4***	0.9462***
FTSE	-0.0004	0.0102	-0.1327	3.7329	1901.21***	0.9539***
HSI	-0.0003	0.0108	0.0833	12.9540	20546***	0.9266***
IBEX	-0.0006	0.0134	0.0471	5.3487	3901.44***	0.9516***
KOSPI	-0.0004	0.0133	-0.3305	4.5158	2778.1***	0.9693***
NASDAQ	-0.0004	0.0153	0.1548	6.0339	4935.59***	0.9410***
NIKKEI	-0.0005	0.0120	-0.4113	10.8710	15576.3***	0.9161***
SPX	0.0000	0.0129	-0.1334	6.7134	6094.1***	0.9252***
SSMI	-0.0003	0.0104	-0.0514	6.3998	5552.89***	0.9290***

The volatility process is estimated with a use of a realized variance (volatility) approach. The realized variance makes use of the high-frequency data, which have been made freely available more frequently in past years, yielding consistent and efficient estimates of the true variance process (Barndorff-Nielsen & Shephard 2002a;b; Hansen & Lunde 2006). The realized variance can be simply seen as the uncentered second moment of the high-frequency series during a specific day. In our case, we use the 5 minutes frequency so that the realized variance is defined as

$$\widehat{\sigma^2_{t,RV}} = \sum_{i=1}^{n} r^2_{t,i} \tag{6.1}$$

where $r_{t,i}$ is a return of the i-th 5-minute interval during day t and n is the number of these 5-minute intervals for a given day. In the analysis, we focus

on the logarithmic volatility[2], i.e. the logarithm of the square root of the realized variance, which is standardly done in the literature, mainly due to the distributional properties and non-negativity of the variance process for eventual simulation of the processes. Later in the text, if we refer to the "volatility process", it stands for the logarithmic volatility process unless stated otherwise. Basic descriptive statistics of volatility series for the analyzed indices are given in Table 6.3. We observe that volatility is positively skewed in general, which is, however, expected as it is bounded to non-negative values. There is no such universality for the excess kurtosis and it reaches rather low values (with exception of the Brazilian BVSPA). This is reflected in rather low normality tests statistics. Even though normality is again strongly rejected, the processes are quite close to being normally distributed, which can be observed in Figure D.1-D.7. In these figures, we can also observe non-normality of the returns process as well as the volatility clustering of the returns series which is evidently connected to the time-varying variance, which is also illustrated. The figures also indicate that it makes more sense to model and analyze the logarithmic volatility rather than the raw volatility as the logarithmic series show much less extreme behavior.

Table 6.3: **Descriptive statistics – logarithmic realized volatility.**

	mean	SD	skewness	excess kurtosis	Jarque-Bera	Shapiro-Wilk
AEX	-4.7234	0.5104	0.4295	0.0420	101.867***	0.9878***
BVSPA	-4.3451	0.3831	0.7076	1.8090	696.286***	0.9711***
CAC	-4.6256	0.4947	0.2309	0.0744	30.1489***	0.9958***
DAX	-4.5366	0.5214	0.3080	-0.0170	52.0617***	0.9938***
DJIA	-4.7770	0.5144	0.5651	0.5544	214.075***	0.9822***
EuroSTOXX	-4.5749	0.5106	0.3048	0.3850	71.0681***	0.9936***
FTSE	-4.8951	0.5192	0.3498	-0.0008	66.4364***	0.9908***
HSI	-4.7845	0.4135	0.5445	0.8328	230.096***	0.9840***
IBEX	-4.6340	0.5138	-0.0192	-0.3360	15.5955***	0.9900***
KOSPI	-4.5905	0.4572	0.4074	0.2205	95.0384***	0.9903***
NASDAQ	-4.6768	0.5177	0.4294	-0.0767	100.512***	0.9851***
NIKKEI	-4.7302	0.4229	0.3049	0.6175	98.6768***	0.9927***
SPX	-4.7562	0.5191	0.5084	0.3827	159.363***	0.9851***
SSMI	-4.8864	0.4720	0.7982	0.4888	377.81***	0.9290***

[2]We analyze the volatility series rather than the changes in volatility as in e.g. Aït-Sahalia *et al.* (2013) due to an assumption of uncorrelated returns series (i.e. $H_x = 0.5$) and persistent volatility series (i.e. $H_y > 0.5$) which brings us to possible cross-persistence of the processes. However, assuming that the changes in volatility are anti-persistent (i.e. $H_y < 0.5$), the analysis would be led to cross-anti-persistence analysis which is not a part of this text. As it turns out, these assumptions are found to be true for the analyzed series.

The realized variance and volatility are also used to standardize the returns series obtaining the r_t^* series. The descriptive statistics summarized in Table 6.4 show that by controlling for the time-varying volatility, the returns series become approximately symmetric with no fat tails. This is again reflected in low normality test statistics. Even though normality is still rejected for all series, by looking at Figure D.1-D.7, we can see that the deviation from normality is very minor and the standardized returns are approximately normally distributed.

Table 6.4: **Descriptive statistics – standardized returns.**

	mean	SD	skewness	excess kurtosis	Jarque-Bera	Shapiro-Wilk
AEX	0.0148	1.0187	0.1109	-0.1289	9.0680**	0.9987**
BVSPA	0.0657	1.1927	0.0379	-0.2682	10.2511***	0.9985***
CAC	0.0089	0.9968	0.0969	-0.2776	15.7931***	0.9978***
DAX	0.0408	0.9756	0.0771	-0.3000	15.596***	0.9978***
DJIA	0.0726	1.0360	0.0099	-0.3944	21.0602***	0.9973***
EuroSTOXX	0.0351	1.0265	0.0907	-0.3108	17.7105***	0.9977***
FTSE	-0.0021	1.0187	0.0094	-0.3578	17.4314***	0.9980***
HSI	0.0007	0.9822	0.1024	-0.3589	20.8989***	0.9970***
IBEX	0.0100	1.0421	0.0804	-0.3199	17.4744***	0.9978***
KOSPI	0.0175	1.0395	0.0281	-0.3789	19.5744***	0.9974***
NASDAQ	0.0388	1.1974	-0.0058	-0.4456	26.8684***	0.9966***
NIKKEI	-0.0198	1.0038	0.0313	-0.4372	25.5617***	0.9967***
SPX	0.0618	1.0657	0.0001	-0.4213	23.9602***	0.9962***
SSMI	0.0109	1.0151	0.0260	-0.2372	7.9918**	0.9989**

In Table D.1-D.3, stationarity tests are summarized. We apply the Augmented Dickey-Fuller test (Dickey & Fuller 1979) with a constant (ADF_1), and with a constant and a time trend (ADF_2) with a maximum of 10 lags based on the Akaike information criterion[3], the KPSS test (Kwiatkowski *et al.* 1992) with a constant ($KPSS_1$), and a constant with a time trend ($KPSS_2$) with the same lag selection, and the Phillips-Perron test (Phillips & Perron 1988). For all three analyzed types of series – returns, standardized returns and volatility – we strongly reject a presence of a unit root ($d = 1$). For returns and standardized returns, stationarity is not rejected for majority of processes and thus we take the returns as asymptotically stationary, i.e. $d < 0.5$ ($H < 1$). For volatility processes, we strongly reject stationarity for practically all processes. Together with the rejection of a unit-root process, we infer that the series of volatility

[3]From the standardly used information criteria – additionally the Schwarz and Hannan-Quinn ones – the Akaike information criterion is selected as it penalizes for additional parameters the least. As we are working with possibly long-range correlated series, this is a favorable attribute. Such choice is applied also later in the text.

are non-stationary but still mean reverting, i.e. $0.5 \leq d \leq 1$ ($1 \leq H \leq 1.5$). These claims are supported by the estimated fractional differencing parameters d which are summarized in Table 6.5. Based on the GPH (Geweke & Porter-Hudak 1983) and local Whittle estimators (Robinson 1995a), we can see that the estimates for volatility are above 0.5, oft-times $\widehat{d} > 0.6$. However, on statistical basis, we are very close to the frontier case between stationarity and mean-reverting non-stationarity – $d = 0.5$. For the both specifications of returns, we cannot reject $d = 0$ (no long-term memory) for practically all cases. These two results – no long-term memory in returns and mean reverting long-term memory in volatility – are sufficient for our ability to estimate the bivariate Hurst exponent or the power law coherency as the sum of the two memory parameters d is below unity for all cases which is an assumption for several proofs and properties of specific long-range cross-correlation processes presented and discussed in Chapter 2 and 3. We can thus follow with an analysis of long-range cross-correlations between returns and volatility.

Table 6.5: **Estimated d parameter for returns and volatility.** The estimates of d are based on the local Whittle and GPH estimators with asymptotic standard errors. Results are given for raw returns (*left*), standardized returns (*middle*), and logarithmic realized volatility (*right*).

	$\widehat{d_{GPH}}$	SE	$\widehat{d_{LW}}$	SE	$\widehat{d_{GPH}}$	SE	$\widehat{d_{LW}}$	SE	$\widehat{d_{GPH}}$	SE	$\widehat{d_{LW}}$	SE
AEX	0.0332	0.0590	0.0297	0.0442	0.0545	0.0536	0.0531	0.0442	0.6206	0.0624	0.6574	0.0442
BVSPA	0.1209	0.0584	0.0536	0.0447	0.0960	0.0618	0.0980	0.0447	0.5575	0.0590	0.5459	0.0447
CAC	0.0588	0.0552	0.0417	0.0442	0.0542	0.0445	0.0485	0.0442	0.6308	0.0565	0.6254	0.0442
DAX	0.0967	0.0613	0.0486	0.0444	0.0781	0.0627	0.0980	0.0444	0.6749	0.0581	0.6706	0.0444
DJIA	-0.0315	0.0544	-0.0457	0.0445	0.1059	0.0640	0.0675	0.0445	0.6972	0.0676	0.6494	0.0445
EuroSTOXX	0.0242	0.0438	0.0251	0.0444	0.0715	0.0589	0.0643	0.0444	0.6423	0.0615	0.6223	0.0444
FTSE	-0.0566	0.0559	-0.0406	0.0444	-0.1206	0.0575	-0.0751	0.0444	0.6800	0.0730	0.6564	0.0444
HIS	-0.0101	0.0623	-0.0187	0.0458	-0.0230	0.0679	0.0573	0.0458	0.6584	0.0556	0.6591	0.0458
IBEX	0.0507	0.0577	0.0399	0.0444	0.1025	0.0688	0.0809	0.0444	0.6353	0.0609	0.6142	0.0444
KOSPI	-0.0587	0.0600	-0.0183	0.0447	0.0176	0.0585	0.0260	0.0447	0.6979	0.0607	0.6619	0.0447
NASDAQ	0.0703	0.0483	0.0383	0.0445	0.1286	0.0611	0.1024	0.0445	0.6664	0.0550	0.6541	0.0445
NIKKEI	-0.0844	0.0582	-0.1220	0.0449	-0.0431	0.0609	-0.0572	0.0449	0.5991	0.0656	0.6304	0.0449
SPX	-0.0037	0.0493	0.0039	0.0445	0.1351	0.0659	0.1116	0.0445	0.6739	0.0548	0.6546	0.0445
SSMI	0.0547	0.0654	0.0208	0.0444	0.1400	0.0586	0.0905	0.0444	0.7447	0.0731	0.7024	0.0444

6.2 Results

Before estimating the bivariate Hurst exponent and the power law coherency, we need to test whether the processes are long-range cross-correlated at all. By not doing so, we might arrive at misleading results. After testing the cross-persistence of the processes, we proceed with estimating and discussing the potential power law coherency. Only after that it makes sense to estimate the bivariate Hurst exponent if the power law coherency is found. Even if the power law coherency is not found, it does not make the analysis worthless as it leads to a particular specification of the bivariate data-generating process which is discussed in the following section.

6.2.1 Testing for power-law cross-correlations

The leverage effect is usually associated with a non-zero correlation between returns and volatility (sometimes the increment of volatility) or their cross-correlation at short lags. The initial step of the analysis is thus an investigation of the cross-correlation function. In Figure 6.1, we show the cross-correlation functions between the standardized returns and volatility for all analyzed indices as well as their average cross-correlation for a specific lag (with 99% confidence intervals). The functions are separated according to the geographical region – the black shade lines are attributed to the European and North American indices and the grey shade ones to the others.

For all indices, we observe strongly asymmetric cross-correlation function – the past values of the standardized returns influence the present values of volatility. Moreover, the standardized returns and volatility are not only cross-correlated but also correlated. On contrary, the past values of volatility do not influence the current values of returns as the cross-correlations vanish to a noise level. Interestingly, the cross-correlations for negative lags (from returns to volatility) decay (approach zero from below) very slowly. This is nicely reflected in the Granger-causality testing which shows the standardized returns Granger-cause volatility with the exception of Hong Kong HSI. The reverse Granger-causality is not found in any of the analyzed indices at at least 5% significance level (Table 6.6). The testing statistics are based on the VAR estimation with HAC standard errors and Akaike information criterion selected lags with a maximum of 20. The procedure is feasible even for our case of the borderline mean-reverting non-stationarity (Bauer & Maynard 2012).

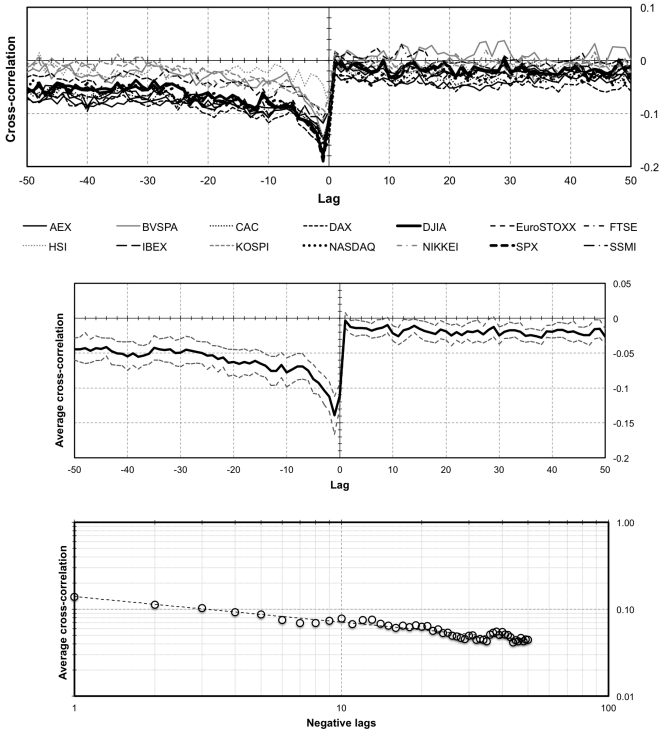


Figure 6.1: **Cross-correlation functions for standardized returns and volatility.** Cross-correlation functions for all analyzed indices (*top*) and with averaged cross-correlations and respective 99% confidence intervals of the mean (*middle*) are shown for the standardized returns r_t^* and logarithmic realized volatility with lags of the standardized returns. The power-law scaling of the cross-correlations (in absolute values) for negative lags (in absolute values) is shown in the log-log scale (*bottom*).

	returns → volatility F-statistic	p-value	volatility → returns F-statistic	p-value	VAR lags *(AIC-based)*
AEX	14.2750	0.0000	0.7644	0.7594	10
BVSPA	8.3519	0.0000	0.7463	0.6811	10
CAC	15.8670	0.0000	1.2393	0.2601	10
DAX	17.2540	0.0000	1.5840	0.1049	10
DJIA	25.8010	0.0000	1.6977	0.0677	11
EuroSTOXX	21.0620	0.0000	0.8577	0.5819	11
FTSE	9.7625	0.0000	1.6630	0.0513	15
HSI	1.2826	0.2277	1.3124	0.2105	11
IBEX	12.5120	0.0000	0.9003	0.5319	10
KOSPI	4.6681	0.0000	0.8743	0.5652	11
NASDAQ	12.8370	0.0000	1.5384	0.1108	11
NIKKEI	10.3330	0.0000	1.1824	0.3015	9
SPX	25.9150	0.0000	1.4751	0.1336	11
SSMI	13.5970	0.0000	0.9782	0.4601	10

Table 6.6: **Granger causality test for leverage effect.** The null hypothesis of 'In $x \to y$, x does not Granger-cause y' using VAR specification with number of lags based on Akaike information criterion (AIC) with a maximum of 20 lags. p-values are based on heteroskedasticity and auto-correlation consistent (HAC) standard error.

The averaged values of the cross-correlation coefficients only emphasize the observations noted earlier. However, we observe that the levels of the cross-correlations differ for the specific indices and we observe that the European and the US indices possess cross-correlations further from the zero values compared to the Asian and Brazilian ones, which poses a question whether all or at least some of the indices can be characterized as the ones with long-range cross-correlated returns and volatility. As the first hint to see whether the cross-correlations are indeed the long-range ones, we also show the averaged cross-correlation function for the negative lags in a double logarithmic scale in Figure 6.1. At the first glance, the power-law fit seems to be a plausible approximation of the decay of cross-correlations so that the long-range cross-correlations seem to be present. Practically the same results are obtained for the relationship between the returns and volatility rather than the standardized returns and volatility. We follow the analysis with the standardized returns only as the results are practically the same for both cases and the standardized returns are more suitable for the potential modeling, which we present at the end of this chapter. Nevertheless, an inspection of the cross-correlation function is only the first hint of the presence of the long-range cross-correlations and it is followed by a statistical analysis and testing.

Index	ξ_5	p-value	ξ_{10}	p-value	ξ_{25}	p-value	ξ_{50}	p-value
AEX	0.7971	0.0180	1.3769	0.0010	2.9179	0.0000	4.9673	0.0000
BVSPA	0.6324	0.0440	0.8539	0.0050	1.6169	0.0020	2.4879	0.0050
CAC	0.8034	0.0110	1.3913	0.0010	3.1468	0.0000	5.5886	0.0000
DAX	1.0206	0.0030	1.7253	0.0020	3.8558	0.0000	6.9938	0.0000
DJIA	0.8115	0.0040	1.3270	0.0010	2.5972	0.0000	4.3642	0.0000
EuroSTOXX	0.7642	0.0120	1.2350	0.0010	2.8251	0.0000	4.9782	0.0000
FTSE	0.5489	0.0240	0.8808	0.0130	1.6708	0.0060	2.6444	0.0020
HSI	0.3389	0.1109	0.4836	0.4126	1.0182	0.1409	1.8444	0.1229
IBEX	0.9093	0.0160	1.5823	0.0000	3.5779	0.0000	6.4400	0.0000
KOSPI	0.5590	0.0210	0.8094	0.0819	1.3578	0.0470	1.8767	0.2368
NASDAQ	0.9395	0.0000	1.5137	0.0000	3.0030	0.0000	5.0918	0.0000
NIKKEI	0.5202	0.0440	0.8323	0.0060	1.5507	0.0060	2.2920	0.0190
SPX	0.8517	0.0000	1.3762	0.0000	2.7585	0.0000	4.5743	0.0000
SSMI	0.6525	0.0080	1.1695	0.0130	2.7325	0.0000	5.0963	0.0000

Table 6.7: **Aggregate cross-correlations test for leverage effect.** Testing statistics ξ_k are calculated for lags $k = 5, 10, 25, 50$. p-values are based on moving-block bootstrapping with a block size of 25 and 1,000 repetitions.

Presence of long-range cross-correlations between the standardized returns and volatility is tested with a use of the two tests introduced in this text – aggregate cross-correlation test (ACC) and rescaled covariance test (RCT), which have been shown to possess the most plausible statistical properties. For ACC, we use the specification with lags $k = 5, 10, 25, 50$ and the moving-block bootstrap with blocks of 25 and 1,000 bootstrapped repetitions. The results are summarized in Table 6.7 and these are very straightforward – apart from Hong Kong HSI and South Korean KOSPI, the processes are long-range cross-correlated and the evidence is statistically very strong. For RCT, we use $q = 5, 10, 30, 50$ and we set $d_x = 0$ for the standardized returns and as an estimate of d_y, we take the average value of the GPH and local Whittle estimators (Table 6.5). The resulting testing statistics are summarized in Table 6.8. Here, we reject that there are no long-range cross-correlations for all indices. However, we need to keep in mind that the rescaled covariance test does not perform flawlessly for the processes which are weakly correlated. For the two specific cases of HSI and KOSPI, which have been found not to be long-range cross-correlated by the previous test, the correlation coefficients between standardized returns and volatility are quite low (-0.0844 and -0.1043) but also the cross-correlations decay very rapidly, which is evident from Figure 6.1. This might cause the estimated long-term covariance $\widehat{s_{xy,q}}$ to be very low and the resulting $M_{xy,T}(q)$ to be quite volatile. In effect, we treat all indices as possessing

the long-range cross-correlations between standardized returns and volatility with the exception of HSI and KOSPI where the evidence is not sufficient. HSI and KOSPI are thus not further analyzed[4].

Index	$M_{xy,T}(5)$	$M_{xy,T}(10)$	$M_{xy,T}(30)$	$M_{xy,T}(50)$
AEX	43.7552***	41.8675***	35.3081***	32.7150***
BVSPA	499.7210***	548.2053***	644.3780***	698.2298***
CAC	131.8652***	122.6288***	96.8466***	86.4880***
DAX	186.8505***	184.3294***	155.8494***	142.1365***
DJIA	49.6715***	50.0130***	47.6993***	46.4876***
EuroSTOXX	130.8961***	131.1738***	105.9571***	94.9270***
FTSE	-2.4304***	-2.6165***	-2.6506***	-2.5549***
HSI	1047.4010***	1127.9730***	1071.2660***	965.2895***
IBEX	262.3148***	242.0165***	190.0478***	168.3787***
KOSPI	272.6409***	289.9085***	329.9933***	369.3005***
NASDAQ	94.6869***	94.5181***	88.4241***	84.8803***
NIKKEI	252.8393***	241.8864***	220.5060***	223.1482***
SPX	102.3680***	102.6113***	95.6743***	92.4764***
SSMI	159.4167***	154.9947***	130.6708***	119.8352***

Table 6.8: **Rescaled covariance test for leverage effect.** Testing statistic $M_{xy,T}(q)$ is calculated for $q = 5, 10, 30, 50$. d_x is set to 0 zero for the returns processes and d_y is taken as an average of the estimates in Table 6.5. Significance is based on p-values from moving-block bootstrapping with a block size of 25 and 1,000 repetitions.

6.2.2 Power law coherency testing

The next step in analyzing the long-range cross-correlations is evaluating whether the bivariate Hurst exponent is lower than the average of the separate Hurst exponents. We have shown that this can be achieved by looking for a potential power law in the squared spectrum coherency of the processes. In order to do so, we have developed several estimators – three based on the time domain estimators of the bivariate Hurst exponent (DCCA, HXA and DMCA), three based on the estimates of the Hurst exponents in the frequency domain (APE, XPE and LXW) and two based on the squared spectrum coherence itself (again based on XPE and LXW).

[4]The results presented in the following text do not differ qualitatively even if the HSI and KOSPI indices are included.

Index	DCCA	HXA	DMCA	$\text{APE}_{m=0.1T}$	$\text{APE}_{m=0.4T}$
AEX	0.1768	0.1653	0.0566	0.0981	0.0339
BVSPA	0.1112	0.1117	0.2806	0.0685	0.0091
CAC	0.2846	0.2111	0.0784	0.1210	0.0395
DAX	0.1996	0.1727	0.1038	0.1076	0.0377
DJIA	0.1968	0.1941	0.4636	0.0802	0.0221
EuroSTOXX	0.2698	0.1607	0.1313	0.1149	0.0275
FTSE	0.1086	0.1160	0.1317	0.0751	0.0053
IBEX	0.2273	0.1868	0.0866	0.1154	0.0383
NASDAQ	0.0851	0.1684	0.2306	0.0649	0.0262
NIKKEI	0.1345	0.2417	0.6143	0.0421	-0.0075
SPX	0.1625	0.1981	0.3994	0.0980	0.0268
SSMI	0.3905	0.1839	0.0029	0.1161	0.0423

Table 6.9: **Estimates of power law coherency I.** DCCA estimates are based on $s_{min} = 10$ and $s_{max} = T/5$, HXA estimates are based on the jackknife specification with $\tau_{min} = 1$ and $\tau_{max} = 5, \ldots, 20$, and DMCA estimates use $\kappa_{min} = 3$ and $\kappa_{max} = 21$. APE estimates are shown for $m = 0.1T$ and $m = 0.4T$ with Daniell's window of 21.

Table 6.10: **Estimates of power law coherency II.** Estimates of H_ρ are based on LXW and XPE. For both estimators, different specifications are shown. Bandwidth parameter m has two settings – $m = 0.1T$ and $m = 0.4T$ – and three levels of Daniell's windows are used.

		$\widehat{H_{\rho,XPE}}^*$			$\widehat{H_{\rho,XPE}}$			$\widehat{H_{\rho,LXW}}^*$			$\widehat{H_{\rho,LXW}}$		
	smothing	11	21	51	11	21	51	11	21	51	11	21	51
AEX	$m = 0.1T$	0.0957	0.0844	0.0534	0.2115	0.2098	0.2042	0.0744	0.0743	0.0555	0.0760	0.0730	0.0443
	$m = 0.4T$	0.1105	0.1457	0.1537	0.1864	0.2053	0.2141	0.0818	0.1164	0.1484	0.0935	0.1207	0.1354
BVSPA	$m = 0.1T$	0.0043	-0.0237	-0.0703	0.1312	0.1186	0.0979	0.0001	-0.0182	-0.0509	-0.0172	-0.0351	-0.0787
	$m = 0.4T$	0.0755	0.1069	0.1370	0.1464	0.1625	0.1797	0.0654	0.1018	0.1436	0.0678	0.0985	0.1263
CAC	$m = 0.1T$	0.0945	0.0904	0.0786	0.2213	0.2233	0.2288	0.0748	0.0845	0.0794	0.0764	0.0812	0.0664
	$m = 0.4T$	0.1071	0.1476	0.1771	0.1935	0.2158	0.2361	0.0835	0.1186	0.1637	0.0903	0.1221	0.1534
DAX	$m = 0.1T$	0.0653	0.0683	0.0565	0.1915	0.1953	0.1992	0.0593	0.0657	0.0579	0.0565	0.0602	0.0459
	$m = 0.4T$	0.1415	0.1831	0.1909	0.1982	0.2198	0.2288	0.1067	0.1477	0.1779	0.1059	0.1396	0.1566
DJIA	$m = 0.1T$	0.0216	0.0224	0.0111	0.1693	0.1724	0.1728	0.0190	0.0217	0.0119	0.0220	0.0192	-0.0017
	$m = 0.4T$	0.1357	0.1758	0.1987	0.1841	0.2059	0.2222	0.1099	0.1510	0.1906	0.1052	0.1372	0.1629
EuroSTOXX	$m = 0.1T$	0.0471	0.0464	0.0330	0.1883	0.1908	0.1928	0.0442	0.0455	0.0340	0.0425	0.0424	0.0207
	$m = 0.4T$	0.1526	0.2035	0.2271	0.2083	0.2344	0.2502	0.1232	0.1737	0.2256	0.1188	0.1619	0.1954
FTSE	$m = 0.1T$	0.0812	0.0775	0.0346	0.2244	0.2260	0.2125	0.0796	0.0933	0.0710	0.0805	0.0739	0.0241
	$m = 0.4T$	0.1203	0.1649	0.1793	0.1915	0.2165	0.2305	0.0968	0.1434	0.1876	0.0979	0.1308	0.1446
IBEX	$m = 0.1T$	0.0639	0.0587	0.0498	0.1911	0.1934	0.2031	0.0449	0.0517	0.0501	0.0572	0.0577	0.0405
	$m = 0.4T$	0.1098	0.1451	0.1620	0.1800	0.2003	0.2162	0.0825	0.1208	0.1597	0.0860	0.1146	0.1381
NASDAQ	$m = 0.1T$	0.0339	0.0188	-0.0152	0.1696	0.1658	0.1586	0.0240	0.0166	-0.0098	0.0193	0.0061	-0.0347
	$m = 0.4T$	0.0871	0.1198	0.1434	0.1720	0.1906	0.2087	0.0661	0.0962	0.1269	0.0736	0.0958	0.1104
NIKKEI	$m = 0.1T$	0.0029	0.0200	0.0189	0.1585	0.1700	0.1735	0.0051	0.0180	0.0111	0.0041	0.0140	-0.0013
	$m = 0.4T$	0.0510	0.0780	0.1173	0.1521	0.1673	0.1906	0.0387	0.0665	0.1114	0.0451	0.0696	0.1049
SPX	$m = 0.1T$	0.0325	0.0315	0.0196	0.1763	0.1797	0.1818	0.0261	0.0298	0.0214	0.0297	0.0267	0.0056
	$m = 0.4T$	0.1407	0.1712	0.1800	0.1944	0.2114	0.2209	0.1151	0.1451	0.1816	0.1100	0.1330	0.1513
SSMI	$m = 0.1T$	0.0750	0.0782	0.0468	0.2120	0.2166	0.2081	0.0561	0.0599	0.0451	0.0579	0.0587	0.0289
	$m = 0.4T$	0.1263	0.1565	0.1774	0.1943	0.2114	0.2273	0.0971	0.1352	0.1660	0.0985	0.1244	0.1409

The results for the time domain estimators are summarized in Table 6.9. For DCCA, we use $s_{min} = 10$ and $s_{max} = T/5$ with a step of 10; for HXA, we use the jackknife specification with $\tau_{min} = 1$ and $\tau_{max} = 5, \ldots, 20$; and for DMCA, we use $\kappa_{min} = 3$ and $\kappa_{max} = 21$. Even though the results are rather volatile, we observe that all estimates are positive for all analyzed indices regardless the estimator applied. As the time domain estimators have been all shown to be upward biased even for the power law coherency specification, it is not surprising. However, the estimates presented in the table are mostly around or above 0.2 which indicates that the power law coherency is most likely not present between the processes. The table also shows the results for the APE estimator with two specifications of the bandwidth parameter – $m = 0.1T$ and $m = 0.4T$. For the cross-periodogram estimation, we use Daniell's windows of length 21. Even though the estimates get close to 0 for $m = 0.4T$, all but one remain positive which again indicates no power law coherency.

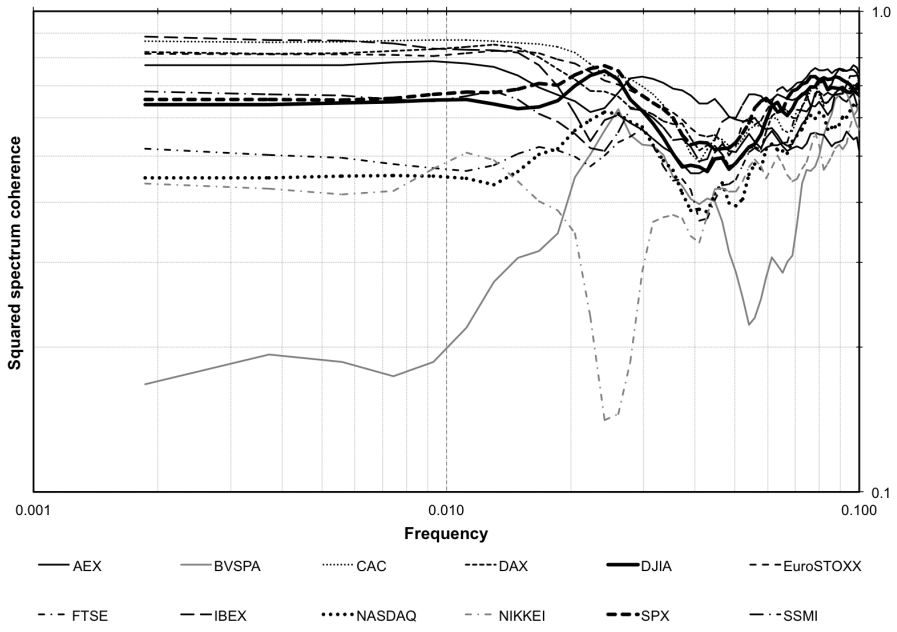


Figure 6.2: **Squared spectrum coherence for leverage effect.** Squared spectrum coherence is estimated using cross-periodogram and periodograms with Daniell's window of 21. Only frequencies $\lambda \leq 0.1$ are showed as the power law coherency might emerge only at the lowest frequencies. Log-log scale is chosen to observe a potential power law.

Table 6.10 summarizes the estimates of H_ρ based on both specifications of LXW and XPE. For estimation of the cross-periodogram, we use three settings of Daniell's windows – 11, 21 and 51 – with two settings of the bandwidth

parameter – $m = 0.1T$ and $m = 0.4T$. The estimates again vary strongly across the indices. However, we again observe that a strong majority of the estimates are higher than zero regardless the level of Daniell's windows or the bandwidth parameter. The power law coherency thus cannot be found even when it is based on these estimators.

Summarizing the results of both types of estimators in a straight forward way – there is no power law coherency between standardized returns and volatility for any of the indices. This is based on the estimates which are either positive or very close to zero. The results are quite similar to the ones of Sela & Hurvich (2012), who analyzed the power law coherency between different monetary aggregates but it was not found and the estimates were also positive in most cases. In our case, the results are more reliable as the series are much longer. Nonetheless, we need to keep in mind that even for the time series length above 3,000 observations, as in our case, the variance of the estimators still remains quite high. The results are supported by the estimated squared spectrum coherencies for the indices based on the cross-periodogram with the length of Daniell's smoothing window of 21 as shown in Figure 6.2. There, we can see that the squared coherence at low frequencies is very stable for most cases and only for Brazilian BVPSA, it follows rather different pattern, which can't, however, be attributed to the power-law scaling.

We thus find no signs of the power law coherency so that there is no need to estimate the bivariate Hurst exponents as these would be equal to the average of 0.5 (the standardized returns with no long-term memory) and the Hurst exponent of the volatility process. Such finding implies that the relationship between standardized returns and logarithmic realized volatility can be possibly characterized as a mixture of short- and long-range cross-correlated processes with correlated innovations. Such possibility is discussed in the next section.

6.3 Discussion

In the previous sections, we have analyzed the leverage effect between the (standardized) returns and volatility of specific stock indices. The results almost universally show that the returns and volatility are long-range cross-correlated and specifically the past values of the returns influence the current levels of volatility. Moreover, these two processes are not only cross-correlated but also correlated. We have also shown that the long-range cross-correlations cannot be attributed to the power law coherency so that the bivariate Hurst exponent is

equal to the average of the two separate Hurst exponents. The relationship between the two processes should not thus be modeled with the mixed-correlated ARFIMA processes, which have been introduced earlier in this text, or with the anti-cointegration processes of Sela & Hurvich (2012). Quite opposite, the behavior of the processes reminds of the processes introduced in Section 3.2 – ARFIMA(0,d,0) and AR(1) processes with correlated innovations for which we also observe strongly asymmetric cross-correlation function where one part shows practically no or very rapidly decaying cross-correlations and the other part shows slowly decaying cross-correlations. If we treat the standardized returns process as the short-memory component of this case and the volatility process as the long-memory component, we can further analyze the leverage effect in this framework.

We combine the empirical findings of this chapter and the standard stylized facts (e.g. in Cont (2001)) to construct a data-generating process for the returns and volatility that mimics the important statistical features discussed in this chapter. We are mainly interested in the leverage effect in its long-range cross-correlated fashion as well as in the non-normality of the returns and the fat tails of the unconditional distribution of returns.

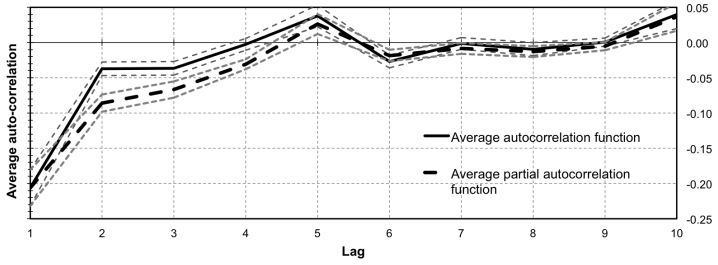


Figure 6.3: **Auto-correlation functions of volatility residuals.** Average auto-correlation and partial auto-correlation functions are showed together with 99% confidence intervals of the means. Residuals $\widehat{\nu}_t$ from the ARFIMA(0,d,0)-filtering form the underlying process of the functions.

We start with the returns specification. The standardized returns are treated as a drift-free serially uncorrelated process with the standard normal distribution while the returns process has a time-varying volatility so that

$$r_t = \varepsilon_t \sigma_t \tag{6.2}$$

$$r_t^* \equiv \varepsilon_t \sim N(0,1). \tag{6.3}$$

This comes from the empirical findings as we have observed that the mean

value of the standardized returns is not statistically different from zero and the variance of the standardized returns is close to a unity for all cases while the distribution is very close to the standard normal one.

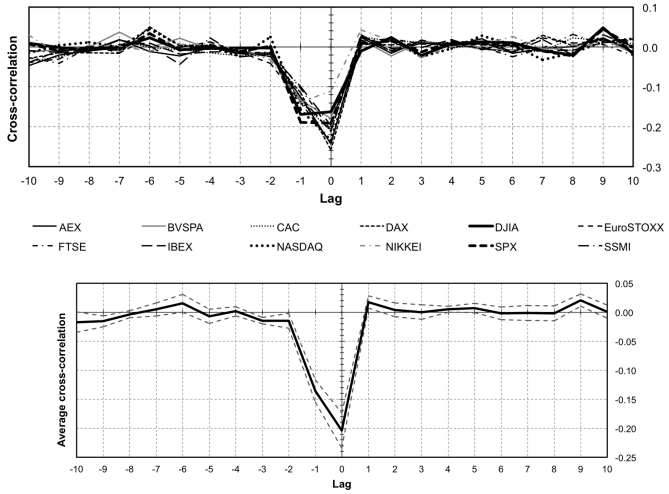


Figure 6.4: **Cross-correlation functions for standardized returns and volatility residuals.** Cross-correlation functions between $\widehat{\varepsilon_t}$ and $\widehat{\nu_t}$ are shown for the separate indices (*top*) and their average value with 99% confidence intervals of the mean (*bottom*).

The specification of the volatility process σ_t comes from our finding that the logarithmic volatility is fractionally integrated (long-range dependent) with parameter d so that

$$\sigma_t = \exp\left(\sum_{k=0}^{\infty} \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} \nu_{t-k}\right) \tag{6.4}$$

$$\nu_t \sim N(0, \sigma_\nu). \tag{6.5}$$

As we have also observed that the standardized returns and volatility processes are not only cross-correlated but also correlated, we need to control for this property. This correlation might easily arise from the correlation between innovations. Therefore, we also assume that

$$\langle \varepsilon_t \nu_t \rangle = \sigma_{\varepsilon\nu}. \tag{6.6}$$

Taking this setting as a starting point for the analysis, we can easily obtain the estimates of ε_t and ν_t as the standardized returns and the fractionally differenced logarithmic volatility process, respectively. Using the average of the estimated d from the GPH and local Whittle estimators, we obtain the

fitted values of ν_t. Now, we can further analyze the statistical properties of $\widehat{\varepsilon_t}$ and $\widehat{\nu_t}$. In Figure 6.3, we illustrate the averaged auto-correlation (ACF) and partial auto-correlation (PACF) functions of the 12 analyzed processes (without HSI and KOSPI) together with the 99% confidence intervals. We can see that ACF shows a significantly negative value at the first lag but falls to insignificant values for higher lags. PACF decays more slowly and shows significant values up to the third lag. Such a behavior is characteristic for the moving average process. In Figure 6.4, we show the cross-correlation functions between $\widehat{\varepsilon_t}$ and $\widehat{\nu_t}$ for separate indices as well as their average value with 99% confidence intervals. We observe that the residuals are negatively correlated and also negatively cross-correlated. However, this negative cross-correlation is significant only for one lag and it implies that $\widehat{\varepsilon_{t-1}}$ influences $\widehat{\nu_t}$. Process ν_t can be thus seen as a vector moving average process. This leads us to the following specification:

$$r_t^* = \varepsilon_t \tag{6.7}$$

$$\nu_t = \psi_t + \delta\psi_{t-1} + \eta\varepsilon_{t-1} \tag{6.8}$$

Assuming that $\langle \varepsilon_s \psi_t \rangle = 0$ for $s \neq t$, this specification leads to the following moments expectations:

$$\sigma_{\varepsilon\nu} \equiv \langle \varepsilon_t \nu_t \rangle = \langle r_t^* \nu_t \rangle = \langle r_t^* \psi_t \rangle = \langle \varepsilon_t \psi_t \rangle \equiv \sigma_{\varepsilon\psi} \tag{6.9}$$

$$\langle r_{t-1}^* \nu_t \rangle = \eta \sigma_\varepsilon^2 \tag{6.10}$$

Parameter η can be thus estimated by using the method of moments using Equation 6.10, i.e.

$$\widehat{\eta} = \frac{\widehat{\mathrm{Cov}(r_{t-1}^* \widehat{\nu_t})}}{\widehat{\sigma}_\varepsilon^2}. \tag{6.11}$$

Parameter δ is then estimated easily through the maximum likelihood estimation of the moving average process. Covariance between disturbances ε_t and ν_t (or ε_t and ψ_t) can be estimated as the covariance between r_t^* and $\widehat{\nu_t}$. Estimates are summarized in Table 6.11. The estimated negative moving average effects δ represents the mean-reverting property of the volatility process and the negative estimate of η shows that there is the short-term memory between innovations of the returns and volatility processes and the positive returns push the volatility downwards and vice versa. The estimated covariance between innovations of the standardized returns and volatility process is negative as well

and it further supports this leverage relationship between variables.

Note that these results are very similar to the ones of Shirota *et al.* (2012) who construct a realized stochastic volatility model which controls for long-term memory of the volatility process as well as for the correlation between innovations of the volatility and returns processes. Even though the model presented in the paper is not the same as the one we present here, several comparisons can be made. First, our specification can be seen as the ARFIMA(0,d,0)-VMA(1) logarithmic volatility process with correlated innovations whereas in Shirota *et al.* (2012), several specifications of ARFIMA(p,d,q) logarithmic volatility models with $p = q = 0$, and $p = 1$ and $q = 0$, and $p = 0$ and $q = 1$ are analyzed. Second, Shirota *et al.* (2012) utilizes Bayesian estimation methods whereas we utilized a multistep procedure combining the maximum likelihood estimation and the method of moments. Third, our results indicate that the correlation between innovations of the volatility and returns processes is close to -0.2 and Shirota *et al.* (2012) get to -0.4. Fourth, long-term memory of the volatility process is characterized by $d \approx 0.6$ which is very close to our average value of $d \approx 0.65$. However, probably the most important finding is that our almost purely empirically-based model has practically the same specification as the state-of-the-art theoretical model of Shirota *et al.* (2012) and just supports that the long-range cross-correlations approach to the process yields meaningful results.

To analyze whether the found empirical properties are sufficient to construct a process which successfully mimics the basic stylized facts and mainly the leverage effect, we use the approximate average values found for the parameters and construct the data-generating process in a following way:

$$\begin{aligned} r_t &= \varepsilon_t \sigma_t \\ \sigma_t &= \exp\left(\sum_{k=0}^{\infty} \frac{\Gamma(k+0.65)}{\Gamma(k+1)\Gamma(0.65)} \nu_{t-k}\right) \\ \nu_t &= \psi_t - 0.3\psi_{t-1} - 0.035\varepsilon_{t-1} \\ \varepsilon_t &\sim N(0,1) \\ \psi_t &\sim N(0,0.25) \\ \rho_{\varepsilon\psi} &= -0.22 \end{aligned} \tag{6.12}$$

As an illustration, we present Figure 6.5 which shows returns and variance processes based on Equation 6.12 with $T = 3000$. We observe that the returns

Index	$\widehat{d}_{avg}$	$\widehat{\delta}$	$\widehat{\eta}$	$\widehat{\sigma_\varepsilon}$	$\widehat{\sigma_\psi}$	$\widehat{\rho_{\varepsilon\nu}}$	$\widehat{\rho_{\varepsilon\psi}}$
AEX	0.6390	-0.2479	-0.0303	1.0187	0.2373	-0.2501	-0.2589
BVSPA	0.5517	-0.1984	-0.0296	1.1927	0.2566	-0.1802	-0.1852
CAC	0.6281	-0.2516	-0.0330	0.9968	0.2317	-0.2196	-0.2280
DAX	0.6728	-0.3029	-0.0299	0.9756	0.2433	-0.2675	-0.2801
DJIA	0.6733	-0.4408	-0.0512	1.0360	0.2848	-0.1891	-0.2083
EuroSTOXX	0.6323	-0.2983	-0.0379	1.0265	0.2663	-0.2527	-0.2648
FTSE	0.6682	-0.3179	-0.0259	1.0187	0.2433	-0.2171	-0.2275
IBEX	0.6248	-0.2443	-0.0285	1.0421	0.2274	-0.2160	-0.2235
NASDAQ	0.6603	-0.2700	-0.0324	1.1974	0.2373	-0.2445	-0.2553
NIKKEI	0.6148	-0.2368	-0.0337	1.0038	0.2360	-0.1112	-0.1152
SPX	0.6643	-0.4073	-0.0529	1.0657	0.2728	-0.2155	-0.2357
SSMI	0.7236	-0.3539	-0.0219	1.0151	0.2109	-0.1994	-0.2109
average	0.6461	-0.2975	-0.0339	1.0491	0.2456	-0.2136	-0.2244

Table 6.11: **Summary of the data-generating process modeling.** The table summarizes estimated parameters of model described between Equation 6.2 and 6.10 for all analyzed indices. Following estimates are shown – long-term memory parameter of the logarithmic volatility processes based on the average of the local Whittle and GPH estimators ($\widehat{d}_{avg}$), vector moving average parameters of the logarithmic volatility residuals ($\widehat{\delta}$ and $\widehat{\eta}$), standard deviations of the standardized returns ($\widehat{\sigma_\varepsilon}$) and innovations of the logarithmic volatility process ($\widehat{\sigma_\psi}$), and correlations between these ($\widehat{\rho_{\varepsilon\nu}}$ and $\widehat{\rho_{\varepsilon\psi}}$).

are evidently not normally distributed and possess fat tails. Also, the clear volatility clustering is illustrated. The simulated variance process also reminds of a standard variance process observed on real financial markets with very long periods of low volatility interspersed with periods of higher volatility or even very extreme volatility as is illustrated at the end of the simulated series. To be able to clearly see the statistical behavior of such simulated processes, we simulate 1,000 processes with $T = 3000$, which is approximately equal to the length of the previously analyzed processes and the specification of Equation 6.12.

The leverage effect is illustrated in Figure 6.6 where we show the average cross-correlation function with 99% confidence intervals together with the log-log scaled left part (negative lags) of the cross-correlation function. The same pattern as for the analyzed indices is observed for the simulated series – the past values of returns have a negative effect on the current values of volatility. The power-law scaling is evident and it is in hand with the long-range cross-correlations detected in the analyzed indices.

Apart from the leverage effect, we are also interested in other stylized facts. To cover at least some of the most basic ones, we present Table 6.12. The

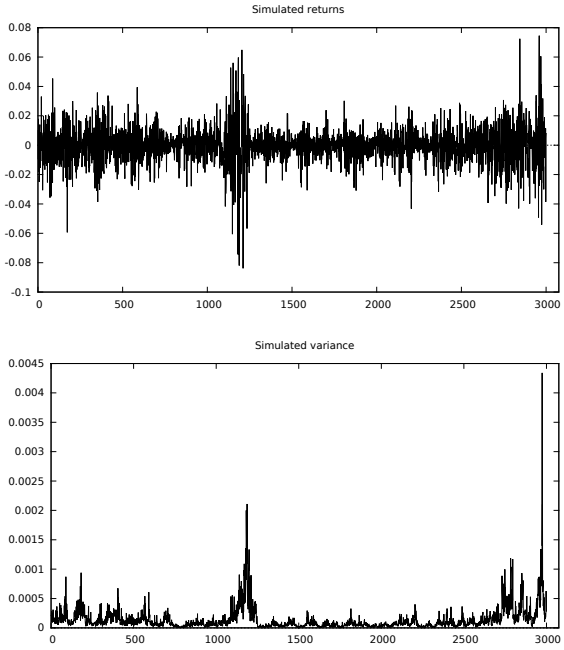


Figure 6.5: **Illustrative realization of a simulated process.** A single realization of a process based on Equation 6.12 with $T = 3000$ – returns (*top*) and variance process (*bottom*).

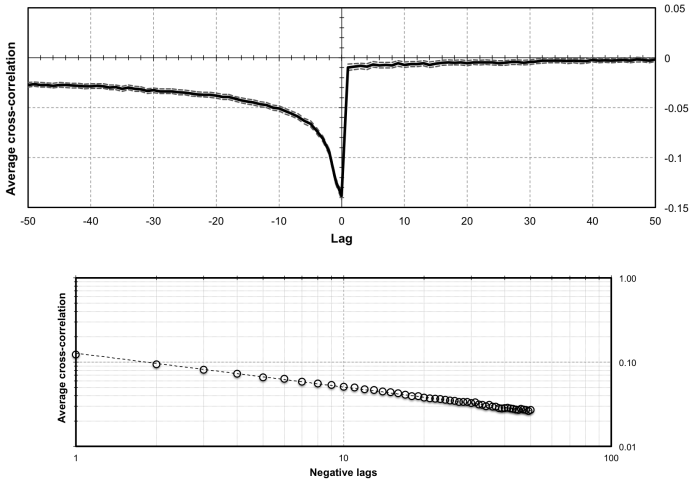


Figure 6.6: **Cross-correlation function of simulated returns and volatility.** Average cross-correlation function with 99% confidence intervals (*top*) and log-log specification of the negative lags of the cross-correlation function (*bottom*) are based on 1,000 simulations of processes based on Equation 6.12 with $T = 3000$.

model is able to produce returns with fat tails, which is, however, nothing new since practically all (G)ARCH-based and stochastic volatility models are able to mimic this property. More importantly, the model is able to produce negatively skewed returns, which implies that extreme negative returns are more frequent than extreme positive returns. And this property is mimicked without any artificial forcing of the property into the model but it comes out of the empirically observed regularities. Interestingly, the negative skewness is not observed for all simulated series as is evident from the positive maximum skewness in the set of simulated series. This is again well in hand with the analyzed stock indices where majority of them was negatively skewed but not all of them.

	average	SD	SD of average	min	max
mean	-0.0006	0.0002	0.0000	-0.0017	0.0001
SD	0.0124	0.0014	0.0000	0.0103	0.0242
skewness	-0.5359	0.3094	0.0098	-2.6578	0.7123
excess kurtosis	5.8091	4.0318	0.1275	1.6383	38.1155

Table 6.12: **Descriptive statistics of simulated series.** Statistics are based on 1,000 simulations of model described in Equation 6.12 with $T = 3000$.

Based on our analysis, the leverage effect is mainly driven by the following factors. First, the innovations of the returns and volatility processes are negatively correlated. Even though the level of this correlation is quite low (around -0.2), this relationship is translated from the innovations into the negative correlation between returns and volatility. This makes perfect sense since the information flow to the financial market is the same for both returns and volatility so that there is no surprise these are correlated. The negative sign of the correlation arises simply from the fact that positive news increase the price of an asset (and thus yields positive return) and simultaneously decrease the tension on the market and in turn also decreases volatility, and inversely, negative news decrease the price of an asset (and brings negative return) and again simultaneously increase nervousness on the market connected with higher volatility. The fact that the correlation is not a perfect one reflects that returns and volatility use different parts of the information flow. Second, the returns process is short-range correlated and the volatility is a long-range cross-correlated process. Connecting this property with the previous point by itself yields long-range cross-correlations between the two processes. The very strong persistence of the logarithmic volatility translates into rather slowly

decaying cross-correlations between the volatility and returns. Third, the innovations of volatility process follow a pattern which can be seen as a vector moving average. The negative sign of the moving average component ensures that the volatility process is mean reverting and thus does not explode. This can be seen as an additional mean reversion property in addition to $d < 1$ for the volatility process. The former can be seen as a short-term mean reversion which ensures that the volatility process does not explode locally. The latter is a long-term mean reversion which guarantees that the process returns to its mean in a long run. On the other hand, the negative effect of the lagged returns innovations implies that the shocks to the returns process have some inertia with respect to the volatility process. Interestingly, the opposite relationship does not hold. However, the effect of shocks in the lagged returns on the volatility vanishes rapidly and it is significant only for one period. Both the negative effects in the vector moving average representation increase the level of cross-correlations but do not affect the speed of decay.

The seemingly strong and persistent causal relationship leading from returns to volatility thus arises from the properties of the data-generating process. The only lag–lead relationship remaining in the process is the one between volatility innovations and standardized returns, which is, however, only short-lived.

Chapter 7

Concluding remarks and future directions

The main aim of this thesis was to introduce a coherent yet still general framework to understand and to treat long-range cross-correlations between two time series. We started from the definition of the long-range cross-correlated processes via the power-law scaling of the cross-correlation function leading to the divergent at origin magnitude of the cross-power spectrum and the power-law scaling of the partial sums of the two processes. We then discussed several specifications of processes yielding the long-range cross-correlations. Probably the most important finding in this area is the fact that long-range cross-correlations simply arise from long-range correlations of the separate processes if these have correlated or cross-correlated innovations. This turns out to be very important for application in economics and finance as most of the processes taking part on the markets gather information from the same information flow. Even though the information content need not be the same for different processes, these are most likely correlated. The long-range cross-correlations can thus be expected to be found between many variables and processes in economics and finance if these are fractionally integrated. This has been shown in the last part of the thesis on the "leverage effect" between returns and volatility, which has been found to arise mainly from the negatively correlated innovations of the separate processes. To be able to show this, we have developed three new tests for presence of long-range cross-correlations between two processes and we have analyzed finite sample properties of the already existing bivariate Hurst exponent estimators together with two new ones, which have been introduced in the text. Apart from the estimators of the bivariate Hurst exponent, the focus has been

put on the power law coherency estimators and as a part of this, several new estimators have been introduced. Even though we have tried to cover the most important aspects of the topic, the thesis does not answer or deal with all the issues connected to the studying of the long-range cross-correlations. We close the dissertation with a short discussion of some of these open or unanswered questions.

First, the definition of long-range cross-correlated processes assumes the bivariate Hurst exponent to be bounded between 0.5 and 1. However, in some cases, it might happen that both processes under study are non-stationary but they are still mean-reverting with Hurst exponents between 1 and 1.5 so that a first-differencing of the series is not desirable. This might easily lead to the bivariate Hurst exponent above 1 which, however, conflicts with our definition of long-range cross-correlations. In practice, this might be the case for variables in both finance (e.g. volatility and traded volume relationship, and realized and implied volatility relationship) and economics (various fractionally integrated macroeconomic series).

Second, the thesis focuses primarily on the case when the bivariate Hurst exponent is between 0.5 and 1 and is thus cross-persistent. However, the theory can be widened to the cross-anti-persistent processes in the same way as in the univariate case, i.e. with the Hurst exponent between 0 and 0.5. Such processes have a power-law decaying cross-correlation function which is, however, absolutely summable and implies zero at origin cross-power spectrum. Introducing the concept of cross-anti-persistence can be helpful to deal with the issue discussed in the previous point as the non-stationary mean-reverting processes which can be turned into stationary anti-persistent processes by first-differencing. These can be in turn approached as cross-anti-persistence.

Third, the issue of consistent and efficient estimation of the cross-power spectrum has not been discussed deeply mainly due to the fact that the thesis focuses on the long-range cross-correlation rather than on the estimation of the cross-power spectrum. Statistical properties of the frequency domain estimators can be improved by using more appropriate estimators of the cross-power spectrum. This leads us to the fourth point which is the utilization of wavelets methodology. Wavelets can help at least partially solve the first and the third point raised earlier. For the former, wavelets can be easily applied on non-stationary series. And for the latter, wavelets can be used to consistently estimate the cross-power spectrum as well as the squared coherency for that matter. Yet again, wavelets are not utilized in this text as this might lead us

away from the main topic of the thesis.

Fifth, the finite sample properties study of all introduced and utilized estimators has been based usually on a very simple setting, namely assuming wide-sense stationarity with normally distributed innovations. However, the statistical properties of the methods can be strongly influenced by a violation of the two concepts. For the former, the statistical properties can be studied under assumptions of time-varying volatility of both series, trends, cycles, periodicity and structural breaks to name the most evident ones which can be present in financial and economic series. For the latter, the statistical performance of the studied tests and estimators can be influenced by various distributions, namely the thickness of the tails, infinite variance and asymmetry.

Last but not least, the application of the proposed methodology is quite wide. In a similar manner as in the case of the leverage effect, there can be other relationships in economics and finance which are standardly treated as causal but can, in fact, be only caused by a mutual effect of the univariate long memory and correlation between innovations. As noted earlier, such effect can be expected in many cases and many pairs of economic variables. It is then sufficient to have at least one process which is long-range dependent to obtain practically spurious causality detected by the standard causality tests.

Study of long-range cross-correlations thus opens a fascinating field with many issues yet to be solved. Moreover, some of the standard stylized facts of economics and finance can be tackled and questioned from the long-range cross-correlations perspective.

Bibliography

Aït-Sahalia, Y., J. Fan, & Y. Li (2013): "The leverage effect puzzle: Disentangling sources of bias at high frequency." *Journal of Financial Economics* **109(1)**: pp. 224–249.

Alessio, E., A. Carbone, G. Castelli, & V. Frappietro (2002): "Second-order moving average and scaling of stochastic time series." *European Physica Journal B* **27**: pp. 197–200.

Alvarez-Ramirez, J., M. Cisneros, C. Ibarra-Valdez, & A. Soriano (2002): "Multifractal Hurst analysis of crude oil prices." *Physica A* **313**: pp. 651–670.

Amblard, P.-O. & J.-F. Coeurjolly (2011): "Identification of the Multivariate Fractional Brownian Motion." *IEEE Transactions on Signal Processing* **59(11)**: pp. 5152–5168.

Amblard, P.-O., J.-F. Coeurjolly, F. Lavancier, & A. Phillipe (2012): "Basic properties of the Multivariate Fractional Brownian Motion." *arXiv:1007.0828* pp. 1–21.

Arianos, S. & A. Carbone (2009): "Cross-correlation of long-range correlated series." *Journal of Statistical Mechanics: Theory and Experiment* **3**: p. P03037.

Baillie, R. & T. Bollerslev (1994): "Cointegration, fractional cointegration, and exchange rate dynamics." *Journal of Finance* **44(2)**: pp. 737–745.

Baillie, R., T. Bollerslev, & H. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **74**: pp. 3–30.

Barabasi, A.-L., P. Szepfalusy, & T. Vicsek (1991): "Multifractal spectra of multi-affine functions." *Physica A* **178**: pp. 17–28.

Barabasi, A. L. & T. Vicsek (1991): "Multifractality of self-affine fractals." *Physical Review A* **44**: pp. 2730–2733.

Barndorff-Nielsen, O. & N. Shephard (2002a): "Econometric analysis of realised volatility and its ise in estimating stochastic volatility models." *Journal of the Royal Statistical Society: Series B* **64**: pp. 253–280.

Barndorff-Nielsen, O. & N. Shephard (2002b): "Estimating quadratic variation using realized variance." *Journal of Applied Econometrics* **17(5)**: pp. 457–477.

Barunik, J., T. Aste, T. Di Matteo, & R. Liu (2012): "Understanding the source of multifractality in financial markets." *Physica A* **391**: pp. 4234–4251.

Barunik, J. & L. Kristoufek (2010): "On Hurst exponent estimation under heavy-tailed distributions." *Physica A* **389(18)**: pp. 3844–3855.

Bauer, D. & A. Maynard (2012): "Persistence-robust surplus-lag Granger causality testing." *Journal of Econometrics* **169**: pp. 293–300.

Bekaert, G. & G. Wu (2000): "Asymmetric volatility and risk in equity markets." *Review of Financial Studies* **13(1)**: pp. 1–42.

Beran, J. (1994): *Statistics for Long-Memory Processes*, volume 61 of *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.

Bertelli, S. & M. Caporin (2002): "A note on calculating autocovariances of long-memory processes." *Journal of Time Series Analysis* **23(5)**: pp. 503–508.

Black, F. (1976): "Studies on stock price volatility changes." *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics* pp. 177–181.

Blahak, U. (2010): "Efficient approximation of the incomplete gamma function for use in cloud model applications." *Geoscientific Model Development* **3**: pp. 329–336.

Bloomfield, P. (2000): *Fourier analysis of time series: An introduction*. John Wiley & Sons, Inc.

Bollerslev, T. & D. Jubinski (1999): "Equity trading volume and volatility: latent information arrivals and common long-run dependencies." *Journal of Business and Economic Statistics* **17**: pp. 9–21.

Bollerslev, T., U. Kretschmer, C. Pigorsch, & G. Tauchen (2009): "A discrete-time model for daily S&P500 returns and realized variations: Jumps and leverage effects." *Journal of Econometrics* **150**: pp. 151–166.

Bollerslev, T., J. Litvinova, & G. Tauchen (2006): "Leverage and volatility feedback effects in high-frequency data." *Journal of Financial Econometrics* **4(3)**: pp. 353–384.

Bollerslev, T. & H. Mikkelsen (1996): "Modeling and pricing long memory in stock market volatility." *Journal of Econometrics* **73**: pp. 151–184.

Bollerslev, T., N. Sizova, & G. Tauchen (2012): "Volatility in equilibrium: Asymmetries and dynamic dependencies." *Review of Finance* **16**: pp. 31–80.

Bouchaud, J.-P., A. Matacz, & M. Potters (2001): "Leverage effect in financial markets: the retarded volatility model." *Physical Review Letters* **87**: p. 228701.

Bouchaud, J.-P. & M. Potters (2001): "More stylized facts of financial markets: leverage effect and downside correlations." *Physica A* **299**: pp. 60–70.

Calvet, L. & A. Fisher (2008): *Multifractal volatility: theory, forecasting, and pricing*. Academic Press.

Cao, G., L. Xu, & J. Cao (2012): "Multifractal detrended cross-correlations between the chinese exchange market and stock market." *Physica A* **391**: pp. 4855–4866.

Carbone, A. (2007): "Algorithm to estimate the Hurst exponent of high-dimensional fractals." *Physical Review E* **76**: p. 056703.

Carr, P. & L. Wu (2009): "Leverage effect, volatility feedback, and self-exciting market disruptions." *Bloomberg Portfolio Research Paper* **3**: pp. 1–66.

Cauchy, A. (1889): *Sur la convergence des séries*. Gauthier-Villars.

Chen, Z., R. Daigler, & A. Parhizgari (2006): "Persistence of volatility in futures markets." *Journal of Futures Markets* **26(6)**: pp. 571–594.

Cheung, Y.-W. & K. S. Lai (1993): "A fractional cointegration analysis of purchasing power parity." *Journal of Business & Economic Statistics* **11(1)**: pp. 103–112.

Coeurjolly, J.-F., P.-O. Amblard, & S. Achard (2012): "Wavelet analysis of the multivariate fractional Brownian motion." *ESAIM: Probability and Statistics* p. doi:10.1051/ps/2012011.

Cont, R. (2001): "Empirical properties of asset returns: stylized facts and statistical issues." *Quantitative Finance* **1(2)(2)**: pp. 223 – 236.

Daniell, P. (1946): "Discussion on symposium on autocorrelation in time series." *Journal of the Royal Statistical Society - Supplement* **8**: pp. 88–90.

Davison, A. & D. Hinkley (1997): *Bootstrap Methds and Their Application*. Cambridge University Press.

Di Matteo, T. (2007): "Multi-scaling in finance." *Quantitatice Finance* **7(1)**: pp. 21–36.

Di Matteo, T., T. Aste, & M. Dacorogna (2003): "Scaling behaviors in differently developed markets." *Physica A* **324**: pp. 183–188.

Di Matteo, T., T. Aste, & M. Dacorogna (2005): "Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development." *Journal of Banking & Finance* **29**: pp. 827–851.

Dickey, D. & W. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* **74**: pp. 427–431.

Dirac, P. (1958): *Principles of quantum mechanics*. Oxford at the Clarendon Press.

Efron, B. (1979): "Bootstrap methods: Another look at the jacknife." *Annals of Statistics* **7**.

Efron, B., R. Tibshirani, & R. Tibshirani (1993): *An introduction to the bootstrap*. Chapman & Hall.

Euler, L. (1738): "Methodus generalis summandi progressiones." *Commentarii Academiae Scientiarum Petropolitanae* **6**: pp. 68–97.

Fama, E. (1965): "The behavior of stock market prices." *Journal of Business* **38**: pp. 34–105.

Fama, E. (1970): "Efficient capital markets: A review of theory and empirical work." *Journal of Finance* **25(2)**: pp. 383–417.

Fleming, J. & C. Kirby (2011): "Long memory in volatility and trading volume." *Journal of Banking & Finance* **35**: pp. 1714–1726.

Forsberg, L. & E. Ghysels (2007): "Why do absolute returns predict volatility so well?" *Journal of Financial Econometrics* **5(1)**: pp. 31–67.

Geweke, J. & S. Porter-Hudak (1983): "The estimation and application of long memory time series models." *Journal of Time Series Analysis* **4(4)**: pp. 221–238.

Gil-Alana, L. & J. Hualde (2009): *Palgrave Handbook of Econometrics 2: Applied Econometrics*, chapter Fractional Integration and Cointegration: An Overview and an Empirical Application. Palgrave.

Giraitis, L., P. Kokoszka, R. Leipus, & G. Teyssière (2003): "Rescaled variance and related tests for long memory in volatility and levels." *Journal of Econometrics* **112**: pp. 265–294.

Giraitis, L., H. Koul, & D. Surgailis (2009): *Large Sample Inference for Long Memory Processes*. Imperial College Press.

Goddard, J. & E. Onali (2012): "Self-affinity in financial asset returns." *Internationa Review of Financial Analysis* **24**: pp. 1–11.

Granger, C. & R. Joyeux (1980): "An introduction to long-memory time series models and fractional differencing." *Journal of Time Series Analysis* **1(1)**: pp. 15–29.

Grech, D. & Z. Mazur (2005): "Statistical properties of old and new techniques in detrended analysis of time series." *Acta Physica Polonica B* **36**: pp. 2403–2413.

Hajian, S. & M. Sadegh Movahed (2010): "Multifractal Detrended Cross-Correlation Analysis of sunspot numbers and river flow fluctuations." *Physica A* **21**: pp. 4942–4957.

Hammoudeh, S., Y. Yuan, T. Chiang, & M. Nandha (2010): "Symmetric and asymmetric US sector return volatilities in presence of oil, financial and economic risks." *Energy Policy* **38**: pp. 3922–3932.

Hansen, P. & A. Lunde (2006): "Realized variance and market microstructure noise." *Journal of Business and Economic Statistics* **24**: pp. 127–218.

He, L.-Y. & S.-P. Chen (2011a): "A new approach to quantify power-law cross-correlation and its application to commodity markets." *Physica A* **390**: pp. 3806–3814.

He, L.-Y. & S.-P. Chen (2011b): "Nonlinear bivariate dependency of price–volume relationships in agricultural commodity futures markets: A perspective from Multifractal Detrended Cross-Correlation Analysis." *Physica A* **390**: pp. 297–308.

Hosking, J. (1981): "Fractional differencing." *Biometrika* **68(1)**: pp. 165–176.

Hurst, H. (1951): "Long term storage capacity of reservoirs." *Transactions of the American Society of Engineers* **116**: pp. 770–799.

Hurvich, C. & W. Chen (2000): "An efficient taper for potentially overdifferenced long memory time series." *Journal of Time Series Analysis* **21(2)**: pp. 156–180.

Ismail, E. (2011): "Asymmetric volatility, leverage effect and financial leverage: A stock market and firm-level analysis." *Middle Eastern Finance and Economics* **15**: pp. 164–186.

Jarque, C. & A. Bera (1980): "Efficient tests for normality, homoskedasticity and serial independence of regression residuals." *Economics Letters* **6(3)**: pp. 255–259.

Jarque, C. & A. Bera (1981): "Efficient tests for normality, homoskedasticity and serial independence of regression residuals: Monte Carlo evidence." *Economics Letters* **7(4)**: pp. 313–318.

Jiang, Z.-Q. & W.-X. Zhou (2011): "Multifractal detrending moving average cross-correlation analysis." *Physical Review E* **84**: p. 016106.

Kantelhardt, J. (2009): *Encyclopedia of Complexity and Systems Science*, chapter Fractal and Multifractal Time Series. Springer.

Kantelhardt, J., S. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, & E. Stanley (2002): "Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series." *Physica A* **316(1-4)**: pp. 87–114.

Kokoszka, P. & M. Taqqu (1996): "Parameter estimation for infinite variance fractional ARIMA." *Annals of Statistics* **24(5)**: pp. 1880–1913.

Kristoufek, L. (2010a): "Long-range dependence in returns and volatility of Central European stock indices." *Bulletin of the Czech Econometric Society* **17(27)**: pp. 50–67.

Kristoufek, L. (2010b): "On spurious anti-persistence in the US stock indices." *Chaos, Solitons and Fractals* **43**: pp. 68–78.

Kristoufek, L. (2010c): "Rescaled range analysis and detrended fluctuation analysis: Finite sample properties and confidence intervals." *AUCO Czech Economic Review* **4**: pp. 236–250.

Kristoufek, L. (2011): "Multifractal height cross-correlation analysis: A new method for analyzing long-range cross-correlations." *EPL* **95**: p. 68001.

Kristoufek, L. (2012): "How are rescaled range analyses affected by different memory and distributional properties? a Monte Carlo study." *Physica A* **391**: pp. 4252–4260.

Künsch, H. (1987): "Statistical aspects of self-similar processes." *Proceedings of the First World Congress of the Bernoulli Society* **1**: pp. 67–74.

Kwiatkowski, D., P. Phillips, P. Schmidt, & Y. Shin (1992): "Testing the null of stationarity against alternative of a unit root: How sure are we that the economic time series have a unit root?" *Journal of Econometrics* **54**: pp. 159–178.

Lavancier, F., A. Philippe, & D. Surgailis (2010): "Atwo-sampletest for comparison of longmemoryparameters." *Journal of Multivariate Analysis* **101(9)**: pp. 2118–2136.

Lavancier, F., A. Phillipe, & D. Surgailis (2009): "Covariance function of vector self-similar process." *Statistics and Probability Letters* **79**: pp. 2415–2421.

Li, J. (2011): "Volatility components, leverage effects, and the return-volatility relations." *Journal of Banking & Finance* **35**: pp. 1530–1540.

Lo, A. (1991): "Long-term memory in stock market prices." *Econometrica* **59(5)**: pp. 1279–1313.

Lobato, I. (1997): "Consistency of the average cross-periodigram in long memory time series." *Journal of Time Series Analysis* **18**: pp. 137–155.

Lobato, I. (1999): "A semiparametric two-step estimator in a multivariate long memory model." *Journal of Econo* **90**: pp. 129–153.

MacLaurin, C. (1742): *A Treatuse of Fluxions*. T.W. and T. Ruddimans.

Mandelbrot, B. (1966): "Is there persistence in stock price movements?" In "Seminar on the Analysis of Security Prices," University of Chicago, Graduate School of Business.

Mandelbrot, B. (1967): "The variation of some other speculative prices." *Journal of Business* **40(4)**: pp. 393–413.

Mandelbrot, B. (1971): "Analysis of long-run dependence in economics: The R/S techniques." *Econometrica* **39**: pp. 68–69.

Mandelbrot, B. (1972): "Statistical methodology for non-periodic cycles: from covariance to R/S analysis." *Annals of Economic and Social Measurement* **1(3)**: pp. 259–290.

Mandelbrot, B. (1999): "A Multifractal Walk down Wall Street." *Scientific American* **280**: pp. 70–73.

Mandelbrot, B. (2005): "The inescapable need for fractal tools in finance." *Annals of Finance* **1**.

Mandelbrot, B., A. Fisher, & L. Calvet (1997): "Multifractal model of asset returns." *Cowles Foundation Discussion Paper* **1164**.

Mandelbrot, B. & J. van Ness (1968): "Fractional brownian motions, fractional noises and applications." *SIAM Review* **10(422)**: pp. 422–437.

Mandelbrot, B. & J. Wallis (1968): "Joah, joseph and operational hydrology." *Water Resources Research* **4**: pp. 909–918.

Marquardt, T. (2007): "Multivariate fractionally integrated CARMA processes." *Journal of Multivariate Analysis* **98**: pp. 1705–1725.

Martens, M., D. van Dijk, & M. de Pooter (2009): "Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements." *International Journal of Forecasting* **25**: pp. 282–303.

Martin, V. & N. Wilkins (1999): "Indirect estimation of ARFIMA and VARFIMA models." *Journal of Econometrics* **93(1)**: pp. 149–175.

Nielsen, F. (2011): "Local Whittle estimation of multi-variate fractionally integrated processes." *Journal of Time Series Analysis* **32(3)**: p. 317–335.

Nielsen, M.-O. (2004a): "Efficient inference in multivariate fractionally integrated time series models." *Econometrics Journal* **7**: pp. 63–97.

Nielsen, M.-O. (2004b): "Spectral analysis of fractional cointegrated systems." *Economics Letters* **83**: pp. 225–231.

Pagan, A. (1996): "The econometrics of financial markets." *Journal of Empirical Finance* **3**: pp. 15–102.

Pagan, A. & G. Schwert (1990): "Alternative models for conditional stock volatility." *Journal of Econometrics* **45**: pp. 267–290.

Peng, C., S. Buldyrev, A. Goldberger, S. Havlin, M. Simons, & H. Stanley (1993): "Finite-size effects on long-range correlations: Implications for analyzing DNA sequences." *Physical Review E* **47(5)**: pp. 3730–3733.

Peng, C., S. Buldyrev, S. Havlin, M. Simons, H. Stanley, & A. Goldberger (1994): "Mosaic organization of DNA nucleotides." *Physical Review E* **49(2)**: pp. 1685–1689.

Phillips, P. & P. Perron (1988): "Testing for a Unit Root in Time Series Regression." *Biometrika* **75**: pp. 335–346.

Podobnik, B., I. Grosse, D. Horvatic, S. Ilic, P. C. Ivanov, & H. E. Stanley (2009a): "Quantifying cross-correlations using local and global detrending approaches." *European Physical Journal B* **71**: pp. 243–250.

Podobnik, B., D. Horvatic, A. Lam Ng, H. Stanley, & P. Ivanov (2008): "Modeling long-range cross-correlations in two-component arfima and fiarch processes." *Physica A* **387**: pp. 3954–3959.

Podobnik, B., D. Horvatic, A. Petersen, & H. E. Stanley (2009b): "Cross-correlations between volume change and price change." *PNAS* **106(52)**: pp. 22079–22084.

Podobnik, B., Z.-Q. Jiang, W.-X. Zhou, & H. E. Stanley (2011): "Statistical tests for power-law cross-correlated processes." *Physical Review E* **84**: p. 066118.

Podobnik, B. & H. Stanley (2008): "Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series." *Physical Review Letters* **100**: p. 084102.

Poon, S.-H. & C. Granger (2003): "Forecasting volatility in financial markets: A review." *Journal of Economic Literature* **41**: pp. 478–539.

Poon, S.-H. & C. Granger (2005): "Practical issues in forecasting volatility." *Financial Analysts Journal* **61(1)**: pp. 45–56.

Ravishanker, N. & B. Ray (1997): "Bayesian analysis of vector ARFIMA processes." *Australian Journal of Statistics* **39(3)**: pp. 295–311.

Robinson, P. (1994): "Semiparametric analysis of long-memory time series." *Annals of Statistics* **22**: pp. 515–539.

Robinson, P. (1995a): "Gaussian semiparametric estimation of long range dependence." *The Annals of Statistics* **23(5)**: pp. 1630–1661.

Robinson, P. (1995b): "Log-Periodogram Regression of Time Series with Long Range Dependence." *The Annals of Statistics* **23(3)**: pp. 1048–1072.

Samorodnitsky, G. (2006): "Long range dependence." *Foundation and Trends® in Stochastic Systems* **1(3)**: pp. 163–257.

Samuelson, P. (1965): "Proof that properly anticipated prices fluctuate randomly." *Industrial Management Review* **6**: pp. 41–49.

Sela, R. (2010): *Three Essays in Econometrics: Multivariate Long Memory Time Series and Applying Regression Trees to Longitudinal Data.* Ph.D. thesis, New York University.

Sela, R. & C. Hurvich (2009): "Computationally efficient methods for two multivariate fractionally integrated models." *Journal of Time Series Analysis* **30(6)**: pp. 631–651.

Sela, R. & C. Hurvich (2012): "The average periodogram estimator for a power law in coherency." *Journal of Time Series Analysis* **33**: pp. 340–363.

Shadkhoo, S. & G. Jafari (2009): "Multifractal detrended cross-correlation analysis of temporal and spatial seismic data." *European Physical Journal B* **72(4)**: pp. 679–683.

Shapiro, S. & M. Wilk (1965): "An analysis of variance test for normality (complete samples)." *Biometrika* **52(3-4)**: pp. 591–611.

Shimotsu, K. (2007): "Gaussian semiparametric estimation of multivariate fractionally intergrated processes." *Journal of Econometrics* **137**: pp. 277–310.

Shirota, S., T. Hizu, & Y. Omori (2012): "Realized stochastic volatility with leverage and long memory." *CIRHE Discussion Paper* **F-869**: pp. 1–34.

Smith, V. & T. Yamagata (2011): "Firm level return-volatility analysis using dynamic panels." *Journal of Empirical Finance* **18**: pp. 847–867.

Sowell, F. (1992): "Maximum likelihood estimation of stationary univariate fractionally integrated time series models." *Journal of Econometrics* **53**: pp. 165–188.

Srinivas, V. & K. Srinivasan (2000): "Post-blackening for modeling dependent annual streamflows." *Journal of Hydrology* **230**.

Talpsepp, T. & M. Rieger (2010): "Explaining asymmetric volatility around the world." *Journal of Empirical* **17**: pp. 938–956.

Taqqu, M., W. Teverosky, & W. Willinger (1995): "Estimators for long-range dependence: an empirical study." *Fractals* **3(4)**: pp. 785–798.

Taqqu, M. & V. Teverovsky (1996): "On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance Time Series." In "A Practical Guide To Heavy Tails: Statistical Techniques and Applications," .

Teverovsky, V., M. Taqqu, & W. Willinger (1999): "A critical look at lo's modified r/s statistic." *Journal of Statistical Planning and Inference* **80(1-2)**: pp. 211–227.

Thomakos, D. & T. Wang (2003): "Realized volatility in the futures markets." *Journal of Empirical Finance* **10**: pp. 321–353.

Tsay, W.-J. (2010): "Maximum likelihood estimation of stationary multivariate ARFIMA processes." *Journal of Statistical Computations & Simulations* **7**: pp. 729–745.

Vandewalle, N. & M. Ausloos (1998): "Crossing of two mobile averages: A method for measuring the roughness exponent." *Physical Review E* **58**: pp. 6832–6834.

Velasco, C. (1999): "Non-stationary log-periodogram regression." *Journal of Econommetrics* **91**: pp. 325–71.

Wall, H. (1948): *Analytic Theory of Continued Fractions*. Chelsea.

Wang, J. & M. Yang (2009): "Asymmetric volatility in the foreign exchange markets." *Journal of International Financial Markets, Institutions & Money* **19**: pp. 597–615.

Wei, W. (2006): *Time series analysis: univariate and multivariate methods*. Pearson Education.

Weron, R. (2002): "Estimating long-range dependence: Finite sample properties and confidence intervals." *Physica A* **312**: pp. 285–299.

Zebende, G. (2011): "DCCA cross-correlation coefficient: Quantifying level of cross-correlation." *Physica A* **390**: pp. 614–618.

Zhao, X., P. Shang, A. Lin, & G. Chen (2011): "Multifractal Fourier detrended cross-correlation analysis of traffic signals." *Physica A* **390**: pp. 3670–3678.

Zhou, W.-X. (2008): "Multifractal detrended cross-correlation analysis for two nonstationary signals." *Physical Review E* **77**: p. 066211.

Appendix A

Cross-power spectra for specific processes

In this part of the Appendix, we present the derivation of cross-correlation structure of processes introduced in Chapter 3. We derive the results in detail for the first case of ARFIMA(0,d,0) processes with correlated innovations and for the other two cases, we keep the same procedure and present only the important points of derivation. Therefore, the last two types of processes are presented only shortly.

A.1 ARFIMA(0,d,0) processes with correlated innovations

As ARFIMA(0,d,0) processes are already represented in their MA(∞) form, we can easily find the cross-correlations of the zeroth and first orders:

$$\rho_{xy}(0) = \frac{\langle x_t y_t \rangle}{\sqrt{\langle x_t^2 \rangle \langle y_t^2 \rangle}} = \frac{\langle (a_0(d_1)\varepsilon_t + a_1(d_1)\varepsilon_{t-1} + \ldots)(a_0(d_2)\nu_t + a_1(d_2)\nu_{t-1} + \ldots) \rangle}{\sigma_x \sigma_y} =$$
$$= \frac{1}{\sigma_x \sigma_y} \sum_{k=0}^{\infty} a_k(d_1) a_k(d_2) \langle \varepsilon_t \nu_t \rangle = \frac{\sigma_{\varepsilon\nu}}{\sigma_x \sigma_y} \sum_{k=0}^{\infty} a_k(d_1) a_k(d_2) \quad \text{(A.1)}$$

$$\rho_{xy}(1) = \frac{\langle x_t y_{t-1}\rangle}{\sqrt{\langle x_t^2\rangle\langle y_t^2\rangle}} = \frac{\langle (a_0(d_1)\varepsilon_t + a_1(d_1)\varepsilon_{t-1} + \ldots)(a_0(d_2)\nu_{t-1} + a_1(d_2)\nu_{t-2} + \ldots)\rangle}{\sigma_x\sigma_y} =$$
$$= \frac{1}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_2)\langle\varepsilon_t\nu_t\rangle = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_2) \quad \text{(A.2)}$$

$$\rho_{xy}(-1) = \frac{\langle x_t y_{t+1}\rangle}{\sqrt{\langle x_t^2\rangle\langle y_t^2\rangle}} = \frac{\langle (a_0(d_1)\varepsilon_t + a_1(d_1)\varepsilon_{t-1} + \ldots)(a_0(d_2)\nu_{t+1} + a_1(d_2)\nu_t + \ldots)\rangle}{\sigma_x\sigma_y} =$$
$$= \frac{1}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_2)\langle\varepsilon_t\nu_t\rangle = \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_2) \quad \text{(A.3)}$$

Using the same procedure as in Equation A.1-A.3, the following cross-correlations patten emerges:

$$\begin{aligned}
\rho_{xy}(0) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_k(d_2)\\
\rho_{xy}(1) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_2)\\
\rho_{xy}(2) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+2}(d_1)a_k(d_2)\\
&\vdots\\
\rho_{xy}(-1) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_2)\\
\rho_{xy}(-2) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+2}(d_2)\\
&\vdots
\end{aligned} \quad \text{(A.4)}$$

With this pattern in mind, we can find a cross-power spectrum as

$$f_{xy}(\lambda) = \frac{1}{2\pi}\frac{\sigma_{\varepsilon\nu}\sigma_x\sigma_y}{\sigma_x\sigma_y}\Big[\sum_{k=0}^{\infty} a_k(d_1)a_k(d_2) + \sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_2)\exp(i\lambda) +$$
$$\sum_{k=0}^{\infty} a_k(d_1)a_{k+2}(d_2)\exp(2i\lambda) + \ldots + \sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_2)\exp(-i\lambda) +$$
$$\sum_{k=0}^{\infty} a_{k+2}(d_1)a_k(d_2)\exp(-2i\lambda) + \ldots\Big] = \frac{\sigma_{\varepsilon\nu}}{2\pi}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_2)\exp(i(k-l)\lambda) \quad \text{(A.5)}$$

which can be simplified to an exponential representation as

$$f_{xy}(\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \left(1 - \exp(i\lambda)\right)^{-d_1} \left(1 - \exp(-i\lambda)\right)^{-d_2}. \tag{A.6}$$

A.2 ARFIMA(0,d,0) and AR(1) processes with correlated innovations

In a similar manner as for the previous case, we can easily find the cross-correlation pattern:

$$\begin{aligned}
\rho_{xy}(0) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_k(d_1)\theta^k \\
\rho_{xy}(1) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_{k+1}(d_1)\theta^k \\
\rho_{xy}(2) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_{k+2}(d_1)\theta^k \\
&\vdots \\
\rho_{xy}(-1) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_k(d_1)\theta^{k+1} \\
\rho_{xy}(-2) &= \frac{\sigma_{\varepsilon\nu}}{\sigma_x\sigma_y} \sum_{k=0}^{\infty} a_k(d_1)\theta^{k+2} \\
&\vdots
\end{aligned} \tag{A.7}$$

Cross-power spectrum can be then written as

$$f_{xy}(\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_l(d_1)\theta^k \exp(i(k-l)\lambda). \tag{A.8}$$

and it can again be rewritten in the exponential representation as

$$f_{xy}(\lambda) = \frac{\sigma_{\varepsilon\nu}}{2\pi} \left(1 - \exp(-i\lambda)\right)^{-d_1} \left(1 - \theta\exp(i\lambda)\right)^{-1}. \tag{A.9}$$

A.3 Mixed-correlated ARFIMA processes

Similarly to the previous cases, we obtain the following rather more complicated structure of the cross-correlations:

$$\rho_{xy}(0) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_k(d_3)+$$
$$\frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_k(d_4) + \frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_k(d_3) + \frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_k(d_4)$$

$$\rho_{xy}(1) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_3)+$$
$$\frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_1)a_k(d_4)+\frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_2)a_k(d_3)+\frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+1}(d_2)a_k(d_4)$$

$$\rho_{xy}(2) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+2}(d_1)a_k(d_3)+$$
$$\frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+2}(d_1)a_k(d_4)+\frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+2}(d_2)a_k(d_3)+\frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_{k+2}(d_2)a_k(d_4)$$

$$\vdots$$

$$\rho_{xy}(-1) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_3)+$$
$$\frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+1}(d_4)+\frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_{k+1}(d_3)+\frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_{k+1}(d_4)$$

$$\rho_{xy}(-2) = \frac{\alpha\gamma\sigma_{13}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+2}(d_3)+$$
$$\frac{\alpha\delta\sigma_{14}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_1)a_{k+2}(d_4)+\frac{\beta\gamma\sigma_{23}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_{k+2}(d_3)+\frac{\beta\delta\sigma_{24}}{\sigma_x\sigma_y}\sum_{k=0}^{\infty} a_k(d_2)a_{k+2}(d_4)$$

$$\vdots \qquad \text{(A.10)}$$

The cross-power spectrum $f_{xy}(\lambda)$ is then

$$\begin{aligned} f_{xy}(\lambda) = \frac{\alpha\gamma\sigma_{13}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_3)\exp(i(k-l)\lambda) + \\ \frac{\alpha\delta\sigma_{14}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_1)a_l(d_4)\exp(i(k-l)\lambda) + \\ \frac{\beta\gamma\sigma_{23}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_2)a_l(d_3)\exp(i(k-l)\lambda) + \\ \frac{\beta\delta\sigma_{24}}{2\pi} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_k(d_2)a_l(d_4)\exp(i(k-l)\lambda). \end{aligned} \quad (A.11)$$

or in the exponential form

$$\begin{aligned} f_{xy}(\lambda) = \frac{1}{2\pi}\Big[\alpha\gamma\sigma_{13}\left(1-\exp(i\lambda)\right)^{-d_1}\left(1-\exp(-i\lambda)\right)^{-d_3} + \\ \alpha\delta\sigma_{14}\left(1-\exp(i\lambda)\right)^{-d_1}\left(1-\exp(-i\lambda)\right)^{-d_4} + \\ \beta\gamma\sigma_{23}\left(1-\exp(i\lambda)\right)^{-d_2}\left(1-\exp(-i\lambda)\right)^{-d_3} + \\ \beta\delta\sigma_{24}\left(1-\exp(i\lambda)\right)^{-d_2}\left(1-\exp(-i\lambda)\right)^{-d_4}\Big]. \end{aligned} \quad (A.12)$$

Appendix B

Tables for Chapter 4

In this part of the Appendix, we present the sizes and powers of the tests introduced in Chapter 4 based on Monte Carlo simulations. For each of the tests – DCCA-based test, aggregate cross-correlations test, partial sums covariance divergence test, and rescaled covariance test – we are interested in the probability that the null hypothesis of 'short-range cross-correlations' is rejected if the null is valid (size of the test), and in the probability that the null is rejected in favor of the alternative hypothesis 'long-range cross-correlations' if the alternative is valid (power of the test). We are also interested in the dependence of the power and size on the correlations level between innovations of the processes as well as length of the series and strength of short- and long-range cross-correlations.

Under the null hypothesis, we assume short-range cross-correlations and for this case, we use correlated noise processes with different correlation levels, and AR(1) processes with correlated innovations. In the latter case, we also inspect how the tests behave under changing strength of short-term memory. To do so, we use three specifications with parameter θ equal to 0.1, 0.5 and 0.8 to control for weak, medium and strong short-term memory.

Under the alternative hypothesis, we assume cross-persistent process and for this matter, we use ARFIMA(0,d,0) processes with correlated innovations. To distinguish between weak and strong long-memory, we examine processes with $d = 0.1$ and $d = 0.4$, respectively.

Table B.1: **Size of $\rho_{DCCA}(s)$ statistic I.** Monte-Carlo-based test size for 1,000 replications of processes $x_t = \varepsilon_t$ and $y_t = \nu_t$ with different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.047	0.145	0.241	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.096	0.195	0.277	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.396	0.530	0.606	0.997	0.998	1.000	1.000	1.000	1.000
	$s = 100$	0.571	0.671	0.726	0.987	0.992	0.994	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.119	0.248	0.335	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.184	0.312	0.408	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.554	0.654	0.719	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 200$	0.667	0.741	0.783	0.992	0.996	0.998	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.419	0.537	0.599	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.549	0.643	0.697	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.792	0.847	0.866	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 1000$	0.850	0.886	0.904	0.997	0.997	0.998	1.000	1.000	1.000

Table B.2: **Size of $\rho_{DCCA}(s)$ statistic II.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.070	0.176	0.257	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.105	0.233	0.322	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.428	0.545	0.619	0.997	0.999	1.000	1.000	1.000	1.000
	$s = 100$	0.581	0.670	0.714	0.992	0.992	0.993	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.111	0.228	0.312	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.209	0.360	0.442	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.548	0.648	0.705	0.999	0.999	1.000	1.000	1.000	1.000
	$s = 200$	0.672	0.752	0.790	0.995	0.996	0.996	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.413	0.534	0.600	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.575	0.670	0.721	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.793	0.839	0.862	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 1000$	0.869	0.898	0.917	1.000	1.000	1.000	1.000	1.000	1.000

Table B.3: **Size of $\rho_{DCCA}(s)$ statistic III.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.5$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.094	0.207	0.289	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.213	0.341	0.419	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.486	0.616	0.679	0.997	0.998	0.999	1.000	1.000	1.000
	$s = 100$	0.591	0.679	0.734	0.988	0.989	0.991	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.194	0.345	0.416	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.308	0.427	0.494	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.598	0.687	0.744	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 200$	0.689	0.771	0.800	0.991	0.991	0.992	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.451	0.559	0.635	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.586	0.682	0.732	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.786	0.846	0.864	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 1000$	0.866	0.900	0.914	0.999	0.999	0.999	1.000	1.000	1.000

Table B.4: **Size of $\rho_{DCCA}(s)$ statistic IV.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.8$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.148	0.288	0.370	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.285	0.414	0.484	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.557	0.649	0.701	0.998	0.991	0.995	1.000	1.000	1.000
	$s = 100$	0.630	0.716	0.759	0.974	0.983	0.984	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.298	0.423	0.488	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.419	0.551	0.609	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.639	0.718	0.758	0.998	0.998	0.999	1.000	1.000	1.000
	$s = 200$	0.734	0.793	0.823	0.986	0.992	0.994	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.559	0.660	0.711	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.614	0.705	0.752	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.829	0.865	0.886	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 1000$	0.858	0.888	0.904	0.994	0.995	0.997	1.000	1.000	1.000

Table B.5: **Power of $\rho_{DCCA}(s)$ statistic I.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.049	0.146	0.229	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.112	0.232	0.319	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.479	0.612	0.663	0.999	0.999	0.999	1.000	1.000	1.000
	$s = 100$	0.604	0.695	0.740	0.981	0.987	0.989	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.106	0.215	0.296	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.223	0.369	0.446	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.598	0.690	0.731	0.998	0.998	0.999	1.000	1.000	1.000
	$s = 200$	0.717	0.791	0.822	0.980	0.988	0.992	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.468	0.584	0.646	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.595	0.689	0.734	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.831	0.864	0.884	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 1000$	0.844	0.877	0.907	0.996	0.997	0.998	1.000	1.000	1.000

Table B.6: **Power of $\rho_{DCCA}(s)$ statistic II.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.4$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$s = 5$	0.087	0.195	0.269	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 10$	0.178	0.311	0.398	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 50$	0.592	0.683	0.722	0.990	0.993	0.995	1.000	1.000	1.000
	$s = 100$	0.686	0.771	0.812	0.952	0.966	0.972	1.000	1.000	1.000
$T = 1000$	$s = 10$	0.168	0.283	0.377	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 20$	0.375	0.490	0.567	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.692	0.771	0.808	0.993	0.996	0.996	1.000	1.000	1.000
	$s = 200$	0.782	0.839	0.861	0.963	0.972	0.976	1.000	1.000	1.000
$T = 5000$	$s = 50$	0.580	0.675	0.723	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 100$	0.693	0.766	0.802	1.000	1.000	1.000	1.000	1.000	1.000
	$s = 500$	0.861	0.900	0.915	0.996	0.996	0.996	1.000	1.000	1.000
	$s = 1000$	0.908	0.933	0.944	0.980	0.980	0.987	1.000	1.000	1.000

Table B.7: **Size of ξ_k statistic I.** Monte-Carlo-based test size for 1,000 replications of processes $x_t = \varepsilon_t$ and $y_t = \nu_t$ with different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.008	0.049	0.094	0.006	0.037	0.076	0.010	0.066	0.121
	$k = 10$	0.010	0.045	0.089	0.012	0.054	0.101	0.010	0.057	0.105
	$k = 25$	0.009	0.047	0.102	0.012	0.057	0.105	0.012	0.050	0.100
	$k = 50$	0.010	0.040	0.095	0.009	0.048	0.098	0.010	0.057	0.097
$T = 1000$	$k = 5$	0.010	0.047	0.095	0.012	0.055	0.095	0.010	0.049	0.097
	$k = 10$	0.016	0.054	0.093	0.009	0.048	0.098	0.007	0.042	0.096
	$k = 25$	0.014	0.048	0.093	0.015	0.055	0.104	0.009	0.049	0.100
	$k = 50$	0.008	0.033	0.082	0.009	0.052	0.098	0.007	0.039	0.083
$T = 5000$	$k = 5$	0.014	0.056	0.114	0.020	0.052	0.109	0.010	0.053	0.108
	$k = 10$	0.016	0.055	0.101	0.011	0.058	0.114	0.010	0.054	0.111
	$k = 25$	0.007	0.042	0.090	0.013	0.057	0.097	0.007	0.055	0.108
	$k = 50$	0.008	0.039	0.098	0.010	0.049	0.095	0.011	0.053	0.106

Table B.8: **Size of ξ_k statistic II.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.010	0.047	0.097	0.012	0.066	0.114	0.015	0.062	0.114
	$k = 10$	0.014	0.058	0.098	0.016	0.067	0.125	0.009	0.056	0.110
	$k = 25$	0.011	0.061	0.111	0.012	0.054	0.095	0.009	0.049	0.093
	$k = 50$	0.015	0.056	0.100	0.011	0.045	0.091	0.012	0.051	0.090
$T = 1000$	$k = 5$	0.004	0.039	0.088	0.015	0.080	0.139	0.021	0.078	0.138
	$k = 10$	0.005	0.032	0.092	0.015	0.061	0.126	0.013	0.058	0.104
	$k = 25$	0.009	0.050	0.098	0.017	0.069	0.127	0.013	0.053	0.109
	$k = 50$	0.011	0.052	0.109	0.010	0.044	0.104	0.012	0.051	0.099
$T = 5000$	$k = 5$	0.014	0.068	0.123	0.017	0.088	0.142	0.026	0.111	0.211
	$k = 10$	0.013	0.060	0.117	0.015	0.064	0.127	0.017	0.081	0.159
	$k = 25$	0.012	0.051	0.115	0.012	0.048	0.108	0.018	0.063	0.130
	$k = 50$	0.009	0.050	0.108	0.014	0.057	0.119	0.011	0.050	0.111

Table B.9: **Size of ξ_k statistic III.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.5$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.014	0.051	0.094	0.043	0.163	0.285	0.089	0.268	0.382
	$k = 10$	0.011	0.056	0.112	0.027	0.101	0.219	0.036	0.148	0.246
	$k = 25$	0.013	0.065	0.126	0.017	0.082	0.155	0.017	0.086	0.157
	$k = 50$	0.019	0.077	0.130	0.033	0.091	0.155	0.018	0.089	0.148
$T = 1000$	$k = 5$	0.020	0.077	0.131	0.078	0.232	0.370	0.190	0.440	0.589
	$k = 10$	0.014	0.065	0.125	0.039	0.152	0.246	0.066	0.203	0.318
	$k = 25$	0.016	0.077	0.148	0.023	0.113	0.192	0.031	0.119	0.215
	$k = 50$	0.016	0.074	0.133	0.025	0.118	0.180	0.031	0.108	0.203
$T = 5000$	$k = 5$	0.015	0.065	0.123	0.480	0.720	0.833	0.814	0.948	0.981
	$k = 10$	0.010	0.060	0.128	0.129	0.367	0.510	0.339	0.616	0.761
	$k = 25$	0.018	0.066	0.127	0.037	0.146	0.262	0.066	0.213	0.352
	$k = 50$	0.024	0.072	0.134	0.037	0.127	0.216	0.038	0.158	0.270

Table B.10: **Size of ξ_k statistic IV.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.8$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.033	0.103	0.174	0.232	0.465	0.617	0.647	0.848	0.929
	$k = 10$	0.023	0.097	0.176	0.099	0.289	0.436	0.285	0.560	0.701
	$k = 25$	0.025	0.091	0.160	0.047	0.165	0.292	0.066	0.229	0.353
	$k = 50$	0.038	0.111	0.189	0.058	0.157	0.273	0.055	0.185	0.287
$T = 1000$	$k = 5$	0.032	0.104	0.184	0.506	0.752	0.844	0.958	0.989	0.997
	$k = 10$	0.019	0.084	0.162	0.233	0.488	0.643	0.563	0.810	0.908
	$k = 25$	0.027	0.090	0.175	0.049	0.186	0.316	0.108	0.330	0.504
	$k = 50$	0.039	0.118	0.200	0.060	0.157	0.276	0.089	0.245	0.375
$T = 5000$	$k = 5$	0.040	0.113	0.191	1.000	1.000	1.000	1.000	1.000	1.000
	$k = 10$	0.023	0.094	0.184	0.929	0.984	0.998	1.000	1.000	1.000
	$k = 25$	0.028	0.101	0.176	0.315	0.637	0.797	0.675	0.938	0.981
	$k = 50$	0.040	0.138	0.216	0.196	0.422	0.592	0.318	0.643	0.803

Table B.11: **Power of ξ_k statistic I.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.009	0.059	0.113	0.026	0.079	0.149	0.033	0.097	0.182
	$k = 10$	0.007	0.049	0.091	0.024	0.086	0.148	0.025	0.092	0.153
	$k = 25$	0.009	0.051	0.097	0.013	0.064	0.130	0.021	0.080	0.151
	$k = 50$	0.010	0.053	0.104	0.013	0.053	0.123	0.018	0.071	0.134
$T = 1000$	$k = 5$	0.016	0.060	0.114	0.026	0.105	0.174	0.038	0.151	0.266
	$k = 10$	0.013	0.065	0.116	0.023	0.102	0.171	0.035	0.144	0.228
	$k = 25$	0.013	0.059	0.112	0.029	0.090	0.157	0.048	0.124	0.216
	$k = 50$	0.019	0.061	0.121	0.026	0.085	0.160	0.032	0.097	0.175
$T = 5000$	$k = 5$	0.016	0.056	0.102	0.127	0.339	0.456	0.344	0.585	0.721
	$k = 10$	0.013	0.061	0.121	0.120	0.302	0.434	0.350	0.614	0.736
	$k = 25$	0.011	0.053	0.098	0.123	0.269	0.379	0.330	0.556	0.676
	$k = 50$	0.020	0.065	0.127	0.098	0.210	0.299	0.264	0.461	0.589

Table B.12: **Power of ξ_k statistic II.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.4$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$k = 5$	0.061	0.175	0.284	0.268	0.503	0.626	0.607	0.843	0.915
	$k = 10$	0.079	0.195	0.292	0.257	0.490	0.612	0.559	0.800	0.868
	$k = 25$	0.106	0.242	0.342	0.231	0.418	0.515	0.458	0.683	0.781
	$k = 50$	0.180	0.310	0.409	0.246	0.406	0.511	0.349	0.542	0.670
$T = 1000$	$k = 5$	0.092	0.199	0.300	0.524	0.739	0.821	0.954	0.997	0.997
	$k = 10$	0.123	0.235	0.321	0.498	0.701	0.781	0.925	0.989	0.995
	$k = 25$	0.170	0.290	0.388	0.458	0.620	0.700	0.827	0.944	0.984
	$k = 50$	0.265	0.385	0.456	0.437	0.587	0.665	0.728	0.865	0.925
$T = 5000$	$k = 5$	0.228	0.366	0.477	0.992	0.997	0.997	1.000	1.000	1.000
	$k = 10$	0.260	0.426	0.521	0.982	0.994	0.997	1.000	1.000	1.000
	$k = 25$	0.342	0.481	0.562	0.945	0.966	0.974	1.000	1.000	1.000
	$k = 50$	0.466	0.549	0.607	0.919	0.947	0.960	1.000	1.000	1.000

Table B.13: **Size of γ_n statistic.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with varying $\theta_x = \theta_y = \theta^*$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$\theta^* = 0$	$T = 500$	0.008	0.052	0.109	0.009	0.053	0.091	0.009	0.045	0.097
	$T = 1000$	0.009	0.050	0.099	0.011	0.063	0.112	0.012	0.049	0.094
	$T = 5000$	0.011	0.045	0.090	0.011	0.061	0.113	0.016	0.055	0.093
$\theta^* = 0.1$	$T = 500$	0.008	0.048	0.092	0.015	0.067	0.109	0.018	0.047	0.113
	$T = 1000$	0.009	0.051	0.101	0.012	0.059	0.101	0.014	0.061	0.130
	$T = 5000$	0.008	0.063	0.113	0.014	0.042	0.101	0.012	0.049	0.101
$\theta^* = 0.5$	$T = 500$	0.008	0.055	0.102	0.012	0.055	0.105	0.010	0.049	0.100
	$T = 1000$	0.012	0.054	0.097	0.013	0.059	0.125	0.013	0.051	0.092
	$T = 5000$	0.016	0.051	0.106	0.021	0.06	0.112	0.010	0.047	0.098
$\theta^* = 0.8$	$T = 500$	0.014	0.070	0.125	0.020	0.063	0.114	0.016	0.059	0.110
	$T = 1000$	0.022	0.082	0.148	0.016	0.062	0.113	0.017	0.064	0.112
	$T = 5000$	0.022	0.072	0.146	0.017	0.066	0.127	0.008	0.068	0.113

Table B.14: **Power of γ_n statistic.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with varying $d_x = d_y = d^*$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$d^* = 0.1$	$T = 500$	0.027	0.09	0.165	0.024	0.087	0.136	0.017	0.073	0.135
	$T = 1000$	0.029	0.125	0.199	0.039	0.105	0.173	0.030	0.107	0.170
	$T = 5000$	0.058	0.17	0.242	0.071	0.185	0.259	0.065	0.159	0.235
$d^* = 0.4$	$T = 500$	0.105	0.251	0.333	0.103	0.226	0.321	0.109	0.234	0.324
	$T = 1000$	0.220	0.367	0.439	0.201	0.327	0.410	0.216	0.339	0.417
	$T = 5000$	0.440	0.572	0.637	0.474	0.572	0.637	0.519	0.588	0.653

Table B.15: **Size of $M_{xy,T}(q)$ statistic I.** Monte-Carlo-based test size for 1,000 replications of processes $x_t = \varepsilon_t$ and $y_t = \nu_t$ with different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.011	0.045	0.094	0.007	0.045	0.092	0.011	0.050	0.099
	$q = 5$	0.011	0.054	0.103	0.009	0.042	0.092	0.010	0.050	0.099
	$q = 10$	0.008	0.041	0.078	0.011	0.042	0.090	0.011	0.052	0.102
	$q = 30$	0.008	0.041	0.078	0.011	0.042	0.090	0.011	0.052	0.102
$T = 1000$	$q = 1$	0.006	0.040	0.088	0.011	0.048	0.101	0.014	0.062	0.094
	$q = 5$	0.015	0.052	0.095	0.012	0.052	0.101	0.014	0.060	0.094
	$q = 10$	0.006	0.056	0.096	0.011	0.053	0.100	0.014	0.053	0.095
	$q = 30$	0.006	0.056	0.096	0.011	0.053	0.100	0.014	0.053	0.095
$T = 5000$	$q = 1$	0.015	0.058	0.105	0.014	0.047	0.100	0.012	0.049	0.101
	$q = 5$	0.010	0.054	0.093	0.014	0.048	0.102	0.012	0.050	0.100
	$q = 10$	0.010	0.052	0.092	0.014	0.048	0.098	0.012	0.050	0.099
	$q = 30$	0.010	0.052	0.092	0.014	0.048	0.098	0.012	0.050	0.099

Table B.16: **Size of $M_{xy,T}(q)$ statistic II.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.011	0.049	0.114	0.006	0.045	0.109	0.009	0.036	0.084
	$q = 5$	0.010	0.058	0.106	0.005	0.048	0.104	0.009	0.038	0.082
	$q = 10$	0.014	0.059	0.114	0.006	0.048	0.108	0.007	0.034	0.085
	$q = 30$	0.014	0.059	0.114	0.006	0.048	0.108	0.007	0.034	0.085
$T = 1000$	$q = 1$	0.006	0.037	0.088	0.013	0.061	0.102	0.018	0.049	0.093
	$q = 5$	0.010	0.042	0.080	0.010	0.063	0.104	0.017	0.049	0.087
	$q = 10$	0.008	0.043	0.075	0.010	0.058	0.105	0.018	0.048	0.090
	$q = 30$	0.008	0.043	0.075	0.010	0.058	0.105	0.018	0.048	0.090
$T = 5000$	$q = 1$	0.016	0.054	0.091	0.014	0.054	0.117	0.011	0.050	0.109
	$q = 5$	0.010	0.053	0.105	0.014	0.053	0.114	0.012	0.050	0.110
	$q = 10$	0.009	0.049	0.108	0.014	0.051	0.115	0.012	0.052	0.109
	$q = 30$	0.009	0.049	0.108	0.014	0.051	0.115	0.012	0.052	0.109

Table B.17: **Size of $M_{xy,T}(q)$ statistic III.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.5$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.008	0.052	0.098	0.006	0.044	0.101	0.012	0.046	0.095
	$q = 5$	0.009	0.055	0.096	0.003	0.043	0.095	0.012	0.047	0.084
	$q = 10$	0.006	0.042	0.086	0.005	0.044	0.092	0.009	0.046	0.083
	$q = 30$	0.006	0.042	0.086	0.005	0.044	0.092	0.009	0.046	0.083
$T = 1000$	$q = 1$	0.018	0.066	0.113	0.011	0.057	0.103	0.012	0.049	0.104
	$q = 5$	0.016	0.057	0.098	0.009	0.053	0.099	0.012	0.046	0.096
	$q = 10$	0.009	0.051	0.106	0.008	0.052	0.093	0.012	0.043	0.096
	$q = 30$	0.009	0.051	0.106	0.008	0.052	0.093	0.012	0.043	0.096
$T = 5000$	$q = 1$	0.015	0.053	0.095	0.006	0.047	0.090	0.015	0.053	0.106
	$q = 5$	0.013	0.053	0.098	0.006	0.042	0.083	0.013	0.055	0.107
	$q = 10$	0.005	0.039	0.079	0.005	0.043	0.079	0.012	0.056	0.106
	$q = 30$	0.005	0.039	0.079	0.005	0.043	0.079	0.012	0.056	0.106

Table B.18: **Size of $M_{xy,T}(q)$ statistic IV.** Monte-Carlo-based test size for 1,000 replications of two AR(1) processes with $\theta_x = \theta_y = 0.8$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.010	0.061	0.118	0.019	0.075	0.135	0.010	0.048	0.104
	$q = 5$	0.006	0.051	0.104	0.013	0.063	0.120	0.008	0.048	0.100
	$q = 10$	0.007	0.038	0.089	0.011	0.058	0.116	0.009	0.047	0.094
	$q = 30$	0.007	0.038	0.089	0.011	0.058	0.116	0.009	0.047	0.094
$T = 1000$	$q = 1$	0.018	0.074	0.123	0.020	0.068	0.130	0.014	0.050	0.097
	$q = 5$	0.014	0.057	0.102	0.015	0.059	0.121	0.012	0.045	0.085
	$q = 10$	0.008	0.053	0.096	0.012	0.054	0.110	0.011	0.047	0.083
	$q = 30$	0.008	0.053	0.096	0.012	0.054	0.110	0.011	0.047	0.083
$T = 5000$	$q = 1$	0.022	0.082	0.145	0.017	0.072	0.120	0.022	0.065	0.108
	$q = 5$	0.020	0.071	0.127	0.016	0.064	0.111	0.017	0.054	0.104
	$q = 10$	0.013	0.070	0.120	0.013	0.058	0.104	0.017	0.053	0.102
	$q = 30$	0.013	0.070	0.120	0.013	0.058	0.104	0.017	0.053	0.102

Table B.19: **Power of $M_{xy,T}(q)$ statistic I.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.1$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.021	0.078	0.163	0.018	0.087	0.148	0.029	0.094	0.141
	$q = 5$	0.054	0.158	0.251	0.081	0.184	0.275	0.103	0.196	0.278
	$q = 10$	0.060	0.167	0.267	0.117	0.232	0.343	0.142	0.254	0.344
	$q = 30$	0.060	0.167	0.267	0.117	0.232	0.343	0.142	0.254	0.344
$T = 1000$	$q = 1$	0.029	0.099	0.176	0.030	0.111	0.172	0.023	0.090	0.166
	$q = 5$	0.059	0.149	0.256	0.097	0.205	0.295	0.094	0.215	0.312
	$q = 10$	0.057	0.173	0.285	0.135	0.252	0.349	0.155	0.283	0.369
	$q = 30$	0.057	0.173	0.285	0.135	0.252	0.349	0.155	0.283	0.369
$T = 5000$	$q = 1$	0.070	0.160	0.255	0.091	0.200	0.283	0.090	0.201	0.282
	$q = 5$	0.101	0.221	0.329	0.187	0.320	0.409	0.195	0.342	0.438
	$q = 10$	0.099	0.224	0.338	0.233	0.368	0.466	0.235	0.399	0.500
	$q = 30$	0.099	0.224	0.338	0.233	0.368	0.466	0.235	0.399	0.500

Table B.20: **Power of $M_{xy,T}(q)$ statistic II.** Monte-Carlo-based test power for 1,000 replications of two ARFIMA(0,d,0) processes with $d_x = d_y = 0.4$ and different correlations $\rho_{\varepsilon\nu}$.

		$\rho = 0$			$\rho = 0.5$			$\rho = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$T = 500$	$q = 1$	0.099	0.222	0.321	0.111	0.229	0.318	0.147	0.272	0.356
	$q = 5$	0.556	0.681	0.750	0.649	0.725	0.768	0.734	0.797	0.839
	$q = 10$	0.597	0.741	0.827	0.772	0.830	0.862	0.869	0.904	0.924
	$q = 30$	0.597	0.741	0.827	0.772	0.830	0.862	0.869	0.904	0.924
$T = 1000$	$q = 1$	0.154	0.311	0.397	0.205	0.339	0.421	0.255	0.371	0.464
	$q = 5$	0.563	0.716	0.789	0.697	0.774	0.814	0.747	0.813	0.846
	$q = 10$	0.631	0.787	0.859	0.817	0.867	0.891	0.857	0.893	0.914
	$q = 30$	0.631	0.787	0.859	0.817	0.867	0.891	0.857	0.893	0.914
$T = 5000$	$q = 1$	0.464	0.589	0.645	0.464	0.584	0.636	0.584	0.685	0.737
	$q = 5$	0.748	0.850	0.894	0.823	0.878	0.899	0.892	0.922	0.934
	$q = 10$	0.783	0.894	0.922	0.898	0.922	0.933	0.934	0.958	0.967
	$q = 30$	0.783	0.894	0.922	0.898	0.922	0.933	0.934	0.958	0.967

Appendix C

Tables for Chapter 5

In this part of the Appendix, we present the finite sample properties of time domain – DCCA, HXA and DMCA – and frequency domain – APE, LXW and XPE – estimators of the bivariate Hurst exponent H_{xy} and power law coherency parameter H_ρ. For all the estimators, we are mainly interested in the dependence of bias, variance and mean squared error on parameters of the estimators and correlation level between innovations. For the time domain estimators, we have different parameters for each estimator. For the frequency domain estimators, we a pair of parameters – the bandwidth parameter m and number of Daniell's windows used in the cross-periodogram. For this purpose, we analyze performance of the estimators for three types of processes. First, we use ARFIMA(0,d,0) processes with correlated innovations and control for weak or strong long-term memory, we use $d = 0.1$ and $d = 0.4$. Second, we utilize ARFIMA(0,d,0) and AR(1) processes with correlated innovations to control for short-term memory bias of the estimators. To do so, we fix $d = 0.4$ and analyze cases of different θ parameter of 0.1, 0.5, and 0.8. Third, we simulate mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$ and $d_2 = d_3 = 0.2$ with varying correlation between innovations ε_2 and ε_3. On the same specification, we also inspect the performance of power law coherency estimators.

Table C.1: **Finite sample properties of DCCA I.** DCCA (*top*) and DCCA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	—	—	—	-0.0481	0.1255	0.0181	-0.0271	0.0717	0.0059
	$s_{min} = 20$	—	—	—	-0.0519	0.1798	0.0350	-0.0229	0.0997	0.0105
	$s_{min} = 50$	—	—	—	-0.0717	0.3781	0.1481	-0.0215	0.1953	0.0386
$T = 1000$	$s_{min} = 10$	—	—	—	-0.0434	0.1060	0.0131	-0.0233	0.0558	0.0037
	$s_{min} = 50$	—	—	—	-0.0560	0.2148	0.0493	-0.0215	0.1098	0.0125
	$s_{min} = 100$	—	—	—	-0.0795	0.3817	0.1520	-0.0120	0.1852	0.0347
$T = 5000$	$s_{min} = 10$	—	—	—	-0.0249	0.0816	0.0073	-0.0135	0.0474	0.0024
	$s_{min} = 50$	—	—	—	-0.0275	0.1098	0.0128	-0.0121	0.0642	0.0043
	$s_{min} = 100$	—	—	—	-0.0314	0.1342	0.0190	-0.0130	0.0791	0.0064
$T = 500$	$s_{min} = 10$	-0.0285	0.0523	0.0036	-0.0287	0.0634	0.0048	-0.0266	0.0692	0.0055
	$s_{min} = 20$	-0.0245	0.0736	0.0060	-0.0253	0.0890	0.0086	-0.0223	0.0962	0.0098
	$s_{min} = 50$	-0.0200	0.1455	0.0216	-0.0241	0.1711	0.0299	-0.0206	0.1887	0.0360
$T = 1000$	$s_{min} = 10$	-0.0212	0.0426	0.0023	-0.0249	0.0524	0.0034	-0.0227	0.0538	0.0034
	$s_{min} = 50$	-0.0155	0.0838	0.0073	-0.0240	0.1018	0.0109	-0.0205	0.1061	0.0117
	$s_{min} = 100$	-0.0129	0.1440	0.0209	-0.0294	0.1663	0.0285	-0.0190	0.1786	0.0323
$T = 5000$	$s_{min} = 10$	-0.0113	0.0340	0.0013	-0.0109	0.0415	0.0018	-0.0129	0.0457	0.0023
	$s_{min} = 50$	-0.0098	0.0462	0.0022	-0.0090	0.0559	0.0032	-0.0114	0.0618	0.0040
	$s_{min} = 100$	-0.0103	0.0572	0.0034	-0.0092	0.0684	0.0048	-0.0121	0.0762	0.0060

Table C.2: **Finite sample properties of DCCA II.** DCCA (*top*) and DCCA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	—	—	—	-0.0673	0.1655	0.0319	-0.0458	0.0861	0.0095
	$s_{min} = 20$	—	—	—	-0.0684	0.2407	0.0626	-0.0394	0.1228	0.0166
	$s_{min} = 50$	—	—	—	-0.0938	0.6945	0.4911	-0.0399	0.2530	0.0656
$T = 1000$	$s_{min} = 10$	—	—	—	-0.0603	0.1314	0.0209	-0.0322	0.0750	0.0067
	$s_{min} = 50$	—	—	—	-0.0670	0.2551	0.0696	-0.0331	0.1448	0.0220
	$s_{min} = 100$	—	—	—	-0.0623	0.4987	0.2526	-0.0357	0.2582	0.0680
$T = 5000$	$s_{min} = 10$	—	—	—	-0.0433	0.1055	0.0130	-0.0221	0.0590	0.0040
	$s_{min} = 50$	—	—	—	-0.0511	0.1447	0.0235	-0.0216	0.0789	0.0067
	$s_{min} = 100$	—	—	—	-0.0610	0.1796	0.0360	-0.0243	0.0963	0.0099
$T = 500$	$s_{min} = 10$	-0.0452	0.0666	0.0065	-0.0411	0.0801	0.0081	-0.0439	0.0829	0.0088
	$s_{min} = 20$	-0.0394	0.0935	0.0103	-0.0340	0.1133	0.0140	-0.0376	0.1182	0.0154
	$s_{min} = 50$	-0.0489	0.1961	0.0408	-0.0321	0.2360	0.0567	-0.0371	0.2430	0.0604
$T = 1000$	$s_{min} = 10$	-0.0284	0.0515	0.0035	-0.0317	0.0648	0.0052	-0.0311	0.0717	0.0061
	$s_{min} = 50$	-0.0246	0.1041	0.0114	-0.0234	0.1269	0.0167	-0.0310	0.1384	0.0201
	$s_{min} = 100$	-0.0358	0.1883	0.0367	-0.0229	0.2148	0.0467	-0.0329	0.2475	0.0623
$T = 5000$	$s_{min} = 10$	-0.0141	0.0435	0.0021	-0.0200	0.0510	0.0030	-0.0211	0.0566	0.0036
	$s_{min} = 50$	-0.0120	0.0588	0.0036	-0.0200	0.0691	0.0052	-0.0203	0.0757	0.0061
	$s_{min} = 100$	-0.0134	0.0727	0.0055	-0.0225	0.0848	0.0077	-0.0227	0.0924	0.0091

Table C.3: **Finite sample properties of DCCA III.** DCCA (*top*) and DCCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	—	—	—	-0.0863	0.1669	0.0353	-0.0538	0.0829	0.0098
	$s_{min} = 20$	—	—	—	-0.0944	0.2411	0.0671	-0.0503	0.1172	0.0163
	$s_{min} = 50$	—	—	—	-0.1190	0.5874	0.3592	-0.0450	0.2482	0.0636
$T = 1000$	$s_{min} = 10$	—	—	—	-0.0720	0.1310	0.0224	-0.0435	0.0692	0.0067
	$s_{min} = 50$	—	—	—	-0.0874	0.2683	0.0796	-0.0384	0.1400	0.0211
	$s_{min} = 100$	—	—	—	-0.1016	0.4775	0.2383	-0.0370	0.2413	0.0596
$T = 5000$	$s_{min} = 10$	—	—	—	-0.0502	0.1061	0.0138	-0.0277	0.0554	0.0038
	$s_{min} = 50$	—	—	—	-0.0548	0.1454	0.0241	-0.0263	0.0747	0.0063
	$s_{min} = 100$	—	—	—	-0.0607	0.1813	0.0366	-0.0283	0.0926	0.0094
$T = 500$	$s_{min} = 10$	-0.0172	0.0564	0.0035	-0.0271	0.0656	0.0050	-0.0332	0.0737	0.0065
	$s_{min} = 20$	-0.0167	0.0794	0.0066	-0.0255	0.0902	0.0088	-0.0297	0.1025	0.0114
	$s_{min} = 50$	-0.0182	0.1573	0.0251	-0.0292	0.1780	0.0325	-0.0256	0.2044	0.0424
$T = 1000$	$s_{min} = 10$	-0.0134	0.0450	0.0022	-0.0191	0.0523	0.0031	-0.0266	0.0603	0.0043
	$s_{min} = 50$	-0.0116	0.0874	0.0078	-0.0198	0.1009	0.0106	-0.0220	0.1191	0.0147
	$s_{min} = 100$	-0.0103	0.1533	0.0236	-0.0222	0.1734	0.0306	-0.0200	0.1988	0.0399
$T = 5000$	$s_{min} = 10$	-0.0093	0.0363	0.0014	-0.0134	0.0424	0.0020	-0.0163	0.0472	0.0025
	$s_{min} = 50$	-0.0101	0.0495	0.0026	-0.0136	0.0569	0.0034	-0.0154	0.0635	0.0043
	$s_{min} = 100$	-0.0115	0.0618	0.0040	-0.0145	0.0696	0.0051	-0.0169	0.0784	0.0064

Table C.4: **Finite sample properties of DCCA IV.** DCCA (*top*) and DCCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	—	—	—	-0.0279	0.1599	0.0264	0.0006	0.0833	0.0069
	$s_{min} = 20$	—	—	—	-0.0589	0.2318	0.0573	-0.0211	0.1156	0.0138
	$s_{min} = 50$	—	—	—	-0.1009	0.5401	0.3019	-0.0413	0.2420	0.0603
$T = 1000$	$s_{min} = 10$	—	—	—	-0.0328	0.1249	0.0167	-0.0137	0.0689	0.0049
	$s_{min} = 50$	—	—	—	-0.0703	0.2522	0.0686	-0.0462	0.1357	0.0205
	$s_{min} = 100$	—	—	—	-0.0791	0.4821	0.2386	-0.0574	0.2317	0.0570
$T = 5000$	$s_{min} = 10$	—	—	—	-0.0400	0.1048	0.0126	-0.0211	0.0516	0.0031
	$s_{min} = 50$	—	—	—	-0.0550	0.1436	0.0236	-0.0300	0.0701	0.0058
	$s_{min} = 100$	—	—	—	-0.0634	0.1781	0.0357	-0.0328	0.0870	0.0086
$T = 500$	$s_{min} = 10$	0.0442	0.0591	0.0054	0.0367	0.0698	0.0062	0.0214	0.0750	0.0061
	$s_{min} = 20$	0.0259	0.0834	0.0076	0.0190	0.0978	0.0099	0.0037	0.1030	0.0106
	$s_{min} = 50$	0.0062	0.1712	0.02934	-0.0021	0.1914	0.0366	-0.0110	0.2071	0.0430
$T = 1000$	$s_{min} = 10$	0.0337	0.0470	0.0033	0.0263	0.0545	0.0037	0.0074	0.0618	0.0039
	$s_{min} = 50$	0.0085	0.0919	0.0085	0.0040	0.1059	0.0112	-0.0199	0.1189	0.0145
	$s_{min} = 100$	0.0017	0.1569	0.0246	-0.0030	0.1751	0.0307	-0.0283	0.1935	0.0383
$T = 5000$	$s_{min} = 10$	0.0132	0.0365	0.0015	0.0063	0.0419	0.0018	-0.0052	0.0454	0.0021
	$s_{min} = 50$	0.0044	0.0499	0.0025	-0.0029	0.0565	0.0032	-0.0141	0.0613	0.0040
	$s_{min} = 100$	0.0020	0.0624	0.0039	-0.0058	0.0692	0.0048	-0.0173	0.0755	0.0060

Table C.5: **Finite sample properties of DCCA V.** DCCA (*top*) and DCCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	—	—	—	0.1326	0.1694	0.0463	0.1553	0.0877	0.0318
	$s_{min} = 20$	—	—	—	0.0737	0.2427	0.0643	0.1021	0.1226	0.0255
	$s_{min} = 50$	—	—	—	-0.0219	0.5988	0.3590	0.0207	0.2548	0.0654
$T = 1000$	$s_{min} = 10$	—	—	—	0.0891	0.1226	0.0230	0.1027	0.0685	0.0152
	$s_{min} = 50$	—	—	—	-0.0061	0.2464	0.0607	0.0113	0.1356	0.0185
	$s_{min} = 100$	—	—	—	-0.0470	0.4676	0.2208	-0.0203	0.2327	0.0545
$T = 5000$	$s_{min} = 10$	—	—	—	-0.0033	0.1019	0.0104	0.0177	0.0544	0.0033
	$s_{min} = 50$	—	—	—	-0.0470	0.1408	0.0220	-0.0179	0.0730	0.0056
	$s_{min} = 100$	—	—	—	-0.0635	0.1761	0.0351	-0.0275	0.0890	0.0087
$T = 500$	$s_{min} = 10$	0.1687	0.0608	0.0321	0.1697	0.0758	0.0345	0.1603	0.0812	0.0323
	$s_{min} = 20$	0.1302	0.0857	0.0243	0.1310	0.1056	0.0283	0.1135	0.1134	0.0257
	$s_{min} = 50$	0.0738	0.1810	0.0382	0.0716	0.2104	0.0494	0.0429	0.2295	0.0545
$T = 1000$	$s_{min} = 10$	0.1303	0.0507	0.0196	0.1250	0.0593	0.0191	0.1133	0.0632	0.0168
	$s_{min} = 50$	0.0620	0.1029	0.0144	0.0575	0.1125	0.0160	0.0337	0.1238	0.0165
	$s_{min} = 100$	0.0306	0.1748	0.0315	0.0280	0.1922	0.0377	0.0079	0.2059	0.0424
$T = 5000$	$s_{min} = 10$	0.0559	0.0374	0.0045	0.0467	0.0436	0.0041	0.0349	0.0485	0.0036
	$s_{min} = 50$	0.0273	0.0503	0.0033	0.0162	0.0592	0.0038	0.0029	0.0649	0.0042
	$s_{min} = 100$	0.0182	0.0617	0.0041	0.0064	0.0728	0.0053	-0.0059	0.0787	0.0062

Table C.6: **Finite sample properties of DCCA VI.** $DCCA_{abs}$ estimator for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	0.0933	0.0628	0.0127	0.0853	0.0634	0.0113	0.0809	0.0652	0.0108
	$s_{min} = 20$	0.1076	0.0908	0.0198	0.0976	0.0897	0.0176	0.0943	0.0916	0.0173
	$s_{min} = 50$	0.1248	0.1943	0.0533	0.1064	0.1872	0.0464	0.1035	0.1946	0.0486
$T = 1000$	$s_{min} = 10$	0.1119	0.0515	0.01516	0.1066	0.0532	0.0142	0.1006	0.0563	0.0133
	$s_{min} = 50$	0.1299	0.1025	0.0274	0.1247	0.1066	0.0269	0.1234	0.1094	0.0272
	$s_{min} = 100$	0.1306	0.1823	0.0503	0.1187	0.1914	0.0507	0.1289	0.1094	0.0525
$T = 5000$	$s_{min} = 10$	0.1409	0.0415	0.0216	0.1389	0.0407	0.0210	0.1348	0.0420	0.0199
	$s_{min} = 50$	0.1494	0.0560	0.0255	0.1477	0.0551	0.0249	0.1449	0.0571	0.0243
	$s_{min} = 100$	0.1511	0.0690	0.0276	0.1494	0.0681	0.0270	0.1475	0.0705	0.0267

Table C.7: **Finite sample properties of DCCA VII.** DCCA estimator of H_ρ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$s_{min} = 10$	0.4321	0.2738	0.2617	0.1628	0.2255	0.0774	0.0333	0.1953	0.0393
	$s_{min} = 20$	0.4427	0.3888	0.3471	0.2404	0.3768	0.1998	0.0345	0.2879	0.0841
	$s_{min} = 50$	0.4418	0.9231	1.0472	0.3526	0.9021	0.9312	0.1317	0.7669	0.6054
$T = 1000$	$s_{min} = 10$	0.4346	0.1987	0.2283	0.1660	0.1590	0.0528	0.0506	0.1461	0.0239
	$s_{min} = 50$	0.4686	0.4156	0.3924	0.3141	0.3975	0.2567	0.1174	0.3487	0.1354
	$s_{min} = 100$	0.4669	0.7759	0.8201	0.3886	0.8120	0.8103	0.2226	0.7477	0.6087
$T = 5000$	$s_{min} = 10$	0.4217	0.1277	0.1942	0.2194	0.1245	0.0636	0.1123	0.1120	0.0252
	$s_{min} = 50$	0.4530	0.1794	0.2374	0.2873	0.1892	0.1183	0.1474	0.1658	0.0492
	$s_{min} = 100$	0.4634	0.2260	0.2658	0.3405	0.2450	0.1760	0.1943	0.2269	0.0892

Table C.8: **Finite sample properties of HXA I.** HXA (*top*) and HXA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	—	—	—	-0.0360	0.0721	0.0065	-0.0271	0.0423	0.0025
	$\tau_{max} = 50$	—	—	—	-0.0524	0.0929	0.0114	-0.0357	0.0535	0.0041
	$\tau_{max} = 100$	—	—	—	-0.0732	0.1063	0.0167	-0.0557	0.0713	0.0082
$T = 1000$	$\tau_{max} = 20$	—	—	—	-0.0223	0.0446	0.0025	-0.0201	0.0286	0.0012
	$\tau_{max} = 50$	—	—	—	-0.0314	0.0608	0.0047	-0.0229	0.0352	0.0018
	$\tau_{max} = 100$	—	—	—	-0.0473	0.0802	0.0087	-0.0307	0.0451	0.0030
$T = 5000$	$\tau_{max} = 20$	—	—	—	-0.0147	0.0185	0.0006	-0.0135	0.0124	0.0003
	$\tau_{max} = 50$	—	—	—	-0.0134	0.0235	0.0007	-0.0110	0.0154	0.0004
	$\tau_{max} = 100$	—	—	—	-0.0150	0.0305	0.0012	-0.0104	0.0197	0.0005
$T = 500$	$\tau_{max} = 20$	-0.0234	0.0321	0.0016	-0.0251	0.0377	0.0020	-0.0267	0.0408	0.0024
	$\tau_{max} = 50$	-0.0304	0.0380	0.0024	-0.0318	0.0452	0.0031	-0.0348	0.0511	0.0038
	$\tau_{max} = 100$	-0.0463	0.0465	0.0043	-0.0481	0.0561	0.0055	-0.0532	0.0665	0.0073
$T = 1000$	$\tau_{max} = 20$	-0.0185	0.0218	0.0008	-0.0180	0.0264	0.0010	-0.0200	0.0275	0.0012
	$\tau_{max} = 50$	-0.0193	0.0271	0.0011	-0.0208	0.0324	0.0015	-0.0225	0.0341	0.0017
	$\tau_{max} = 100$	-0.0248	0.0341	0.0018	-0.0281	0.0406	0.0024	-0.0299	0.0433	0.0028
$T = 5000$	$\tau_{max} = 20$	-0.0132	0.0105	0.0003	-0.0137	0.0111	0.0003	-0.0135	0.0121	0.0003
	$\tau_{max} = 50$	-0.0110	0.0124	0.0003	-0.0117	0.0137	0.0003	-0.0110	0.0149	0.0003
	$\tau_{max} = 100$	-0.0107	0.0151	0.0003	-0.0119	0.0174	0.0004	-0.0103	0.0190	0.0005

Table C.9: **Finite sample properties of HXA II.** HXA (*top*) and HXA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	—	—	—	-0.1103	0.0973	0.0216	-0.0865	0.0413	0.0092
	$\tau_{max} = 50$	—	—	—	-0.1370	0.1234	0.0340	-0.1017	0.0525	0.0131
	$\tau_{max} = 100$	—	—	—	-0.1665	0.1333	0.0455	-0.1294	0.0704	0.0217
$T = 1000$	$\tau_{max} = 20$	—	—	—	-0.0816	0.0591	0.0101	-0.0700	0.0303	0.0058
	$\tau_{max} = 50$	—	—	—	-0.0971	0.0826	0.0162	-0.0774	0.0369	0.0074
	$\tau_{max} = 100$	—	—	—	-0.1191	0.1034	0.0249	-0.0911	0.0460	0.0104
$T = 5000$	$\tau_{max} = 20$	—	—	—	-0.0518	0.0281	0.0035	-0.0479	0.0180	0.0026
	$\tau_{max} = 50$	—	—	—	-0.0546	0.0344	0.0042	-0.0488	0.0212	0.0028
	$\tau_{max} = 100$	—	—	—	-0.0612	0.0442	0.0057	-0.0527	0.0251	0.0034
$T = 500$	$\tau_{max} = 20$	-0.0833	0.0335	0.0081	-0.0868	0.0356	0.0088	-0.0858	0.0395	0.0089
	$\tau_{max} = 50$	-0.0957	0.0397	0.0107	-0.1003	0.0435	0.0119	-0.1004	0.0496	0.0125
	$\tau_{max} = 100$	-0.1186	0.0489	0.0164	-0.1240	0.0550	0.0184	-0.1267	0.0649	0.0203
$T = 1000$	$\tau_{max} = 20$	-0.0678	0.0260	0.0053	-0.0700	0.0281	0.0057	-0.0697	0.0293	0.0057
	$\tau_{max} = 50$	-0.0742	0.0307	0.0064	-0.0770	0.0334	0.0070	-0.0769	0.0354	0.0072
	$\tau_{max} = 100$	-0.0865	0.0372	0.0887	-0.0895	0.0409	0.0097	-0.0903	0.0438	0.0101
$T = 5000$	$\tau_{max} = 20$	-0.0473	0.0138	0.0024	-0.0477	0.0160	0.0025	-0.0478	0.0175	0.0026
	$\tau_{max} = 50$	-0.0480	0.0163	0.0026	-0.0485	0.0185	0.0027	-0.0486	0.0205	0.0028
	$\tau_{max} = 100$	-0.0514	0.0193	0.0030	-0.0522	0.0215	0.0032	-0.0524	0.0242	0.0033

Table C.10: **Finite sample properties of HXA III.** HXA (*top*) and HXA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	—	—	—	-0.0635	0.0856	0.0114	-0.0535	0.0440	0.0048
	$\tau_{max} = 50$	—	—	—	-0.0854	0.1112	0.0197	-0.0665	0.0588	0.0079
	$\tau_{max} = 100$	—	—	—	-0.1044	0.1226	0.0259	-0.0929	0.0806	0.0151
$T = 1000$	$\tau_{max} = 20$	—	—	—	-0.0531	0.0602	0.0064	-0.0416	0.0309	0.0027
	$\tau_{max} = 50$	—	—	—	-0.0694	0.0876	0.0125	-0.0455	0.0396	0.0036
	$\tau_{max} = 100$	—	—	—	-0.0885	0.1065	0.00192	-0.0566	0.0531	0.0060
$T = 5000$	$\tau_{max} = 20$	—	—	—	-0.0330	0.0247	0.0017	-0.0308	0.0142	0.0012
	$\tau_{max} = 50$	—	—	—	-0.0324	0.0325	0.0021	-0.0280	0.0182	0.0011
	$\tau_{max} = 100$	—	—	—	-0.0356	0.0441	0.0032	-0.0275	0.0227	0.0013
$T = 500$	$\tau_{max} = 20$	-0.0278	0.0316	0.0018	-0.0309	0.0339	0.0021	-0.0392	0.0377	0.0030
	$\tau_{max} = 50$	-0.0426	0.0386	0.0033	-0.0452	0.0418	0.0038	-0.0531	0.0475	0.0051
	$\tau_{max} = 100$	-0.0643	0.0487	0.0065	-0.0673	0.0544	0.0075	-0.0767	0.0623	0.0098
$T = 1000$	$\tau_{max} = 20$	-0.0190	0.0225	0.0009	-0.0239	0.0254	0.0012	-0.0278	0.0267	0.0015
	$\tau_{max} = 50$	-0.0288	0.0264	0.0015	-0.0337	0.0303	0.0021	-0.0350	0.0331	0.0023
	$\tau_{max} = 100$	-0.0414	0.0326	0.0028	-0.0464	0.0376	0.0036	-0.0466	0.0417	0.0039
$T = 5000$	$\tau_{max} = 20$	-0.0065	0.0115	0.0002	-0.0105	0.0122	0.0003	-0.0162	0.0131	0.0004
	$\tau_{max} = 50$	-0.0120	0.0133	0.0003	-0.0153	0.0147	0.0005	-0.0184	0.0159	0.0006
	$\tau_{max} = 100$	-0.0174	0.0159	0.0006	-0.0202	0.0182	0.0007	-0.0214	0.0194	0.0008

Table C.11: **Finite sample properties of HXA IV.** HXA (*top*) and HXA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	—	—	—	0.0082	0.0837	0.0071	0.0226	0.0403	0.0021
	$\tau_{max} = 50$	—	—	—	-0.0390	0.1099	0.0136	-0.0121	0.0568	0.0034
	$\tau_{max} = 100$	—	—	—	-0.0750	0.1191	0.0198	-0.0505	0.0784	0.0087
$T = 1000$	$\tau_{max} = 20$	—	—	—	0.0230	0.0518	0.0032	0.0312	0.0273	0.0017
	$\tau_{max} = 50$	—	—	—	-0.0162	0.0801	0.0067	0.0029	0.0374	0.0014
	$\tau_{max} = 100$	—	—	—	-0.0503	0.1013	0.0128	-0.0232	0.0514	0.0032
$T = 5000$	$\tau_{max} = 20$	—	—	—	0.0394	0.0216	0.0020	0.0409	0.0124	0.0018
	$\tau_{max} = 50$	—	—	—	0.0161	0.0300	0.0012	0.0195	0.0166	0.0007
	$\tau_{max} = 100$	—	—	—	-0.0015	0.0414	0.0017	0.0042	0.0216	0.0005
$T = 500$	$\tau_{max} = 20$	0.0515	0.0304	0.0036	0.0452	0.0333	0.0032	0.0368	0.0365	0.0023
	$\tau_{max} = 50$	0.0167	0.0366	0.0016	0.0111	0.0422	0.0019	0.0035	0.0481	0.0023
	$\tau_{max} = 100$	-0.0188	0.0469	0.0026	-0.0253	0.0540	0.0036	-0.0322	0.0630	0.0050
$T = 1000$	$\tau_{max} = 20$	0.0604	0.0222	0.0041	0.0538	0.0246	0.0035	0.0459	0.0250	0.0027
	$\tau_{max} = 50$	0.0306	0.0276	0.0017	0.0247	0.0300	0.0015	0.0174	0.0323	0.0013
	$\tau_{max} = 100$	0.0046	0.0339	0.0012	-0.0018	0.0367	0.0013	-0.0083	0.0419	0.0018
$T = 5000$	$\tau_{max} = 20$	0.0717	0.0119	0.0053	0.0672	0.0121	0.0047	0.0577	0.0121	0.0035
	$\tau_{max} = 50$	0.0464	0.0138	0.0023	0.0425	0.0146	0.0020	0.0346	0.0151	0.0014
	$\tau_{max} = 100$	0.0269	0.0163	0.0010	0.0236	0.0177	0.0009	0.0171	0.0189	0.0006

Table C.12: **Finite sample properties of HXA V.** HXA (*top*) and HXA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	—	—	—	0.1137	0.0687	0.0176	0.1263	0.0320	0.0170
	$\tau_{max} = 50$	—	—	—	0.0538	0.0959	0.0121	0.0780	0.0484	0.0084
	$\tau_{max} = 100$	—	—	—	-0.0019	0.1106	0.0122	0.0213	0.0710	0.0055
$T = 1000$	$\tau_{max} = 20$	—	—	—	0.1290	0.0417	0.0184	0.1350	0.0210	0.0187
	$\tau_{max} = 50$	—	—	—	0.0788	0.0681	0.0108	0.0928	0.0315	0.0096
	$\tau_{max} = 100$	—	—	—	0.0285	0.0916	0.0092	0.0498	0.0457	0.0046
$T = 5000$	$\tau_{max} = 20$	—	—	—	0.1424	0.0156	0.0205	0.1427	0.0098	0.0205
	$\tau_{max} = 50$	—	—	—	0.1052	0.0228	0.0116	0.1064	0.0143	0.0115
	$\tau_{max} = 100$	—	—	—	0.0707	0.0336	0.0061	0.0735	0.0201	0.0058
$T = 500$	$\tau_{max} = 20$	0.1321	0.0271	0.0182	0.1308	0.0295	0.0180	0.1290	0.0307	0.0176
	$\tau_{max} = 50$	0.0921	0.0358	0.0098	0.0881	0.0395	0.0093	0.0857	0.0429	0.0092
	$\tau_{max} = 100$	0.0461	0.0479	0.0044	0.0396	0.0525	0.0043	0.0357	0.0584	0.0047
$T = 1000$	$\tau_{max} = 20$	0.1400	0.0203	0.0200	0.1395	0.0211	0.0199	0.1381	0.0207	0.0195
	$\tau_{max} = 50$	0.1041	0.0256	0.0115	0.1030	0.0274	0.0114	0.0995	0.0295	0.0108
	$\tau_{max} = 100$	0.0675	0.0323	0.0056	0.0657	0.0351	0.0056	0.0602	0.0400	0.0052
$T = 5000$	$\tau_{max} = 20$	0.1533	0.0105	0.0236	0.1513	0.0105	0.0230	0.1480	0.0104	0.0220
	$\tau_{max} = 50$	0.1221	0.0129	0.0151	0.1196	0.0132	0.0145	0.1148	0.0140	0.0134
	$\tau_{max} = 100$	0.0920	0.0159	0.0097	0.0896	0.0166	0.0083	0.0837	0.0185	0.0070

Table C.13: **Finite sample properties of HXA VI.** HXA_{abs} estimator for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	0.0678	0.0365	0.0059	0.0624	0.0354	0.0052	0.0564	0.0363	0.0045
	$\tau_{max} = 50$	0.0615	0.0434	0.0057	0.0566	0.0430	0.0051	0.0504	0.0442	0.0045
	$\tau_{max} = 100$	0.0447	0.0527	0.0048	0.0395	0.0537	0.0044	0.0331	0.0560	0.0042
$T = 1000$	$\tau_{max} = 20$	0.0812	0.0275	0.0074	0.0798	0.0253	0.0070	0.0729	0.0271	0.0060
	$\tau_{max} = 50$	0.0815	0.0320	0.0077	0.0802	0.0302	0.0073	0.0743	0.0320	0.0065
	$\tau_{max} = 100$	0.0744	0.0388	0.0070	0.0740	0.0369	0.0068	0.0685	0.0387	0.0062
$T = 5000$	$\tau_{max} = 20$	0.1033	0.0143	0.0109	0.1012	0.0151	0.0105	0.0946	0.0150	0.0092
	$\tau_{max} = 50$	0.1091	0.0167	0.0122	0.1075	0.0173	0.0119	0.1015	0.0176	0.0106
	$\tau_{max} = 100$	0.1111	0.0197	0.0127	0.1098	0.0201	0.0125	0.1043	0.0207	0.0113

Table C.14: **Finite sample properties of HXA VII.** HXA estimator of H_ρ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\tau_{max} = 20$	0.3797	0.2153	0.1905	0.1446	0.2178	0.0683	0.0556	0.1675	0.0312
	$\tau_{max} = 50$	0.3873	0.1605	0.1758	0.2095	0.1746	0.0744	0.1057	0.1409	0.0310
	$\tau_{max} = 100$	0.4016	0.1324	0.1788	0.2614	0.1442	0.0892	0.1664	0.1222	0.0426
$T = 1000$	$\tau_{max} = 20$	0.3636	0.2209	0.1810	0.0858	0.1977	0.0465	0.0439	0.1586	0.0271
	$\tau_{max} = 50$	0.3715	0.1681	0.1663	0.1563	0.1667	0.0522	0.0730	0.1468	0.0269
	$\tau_{max} = 100$	0.3796	0.1402	0.1638	0.2111	0.1502	0.0671	0.1211	0.1344	0.0327
$T = 5000$	$\tau_{max} = 20$	0.2956	0.2512	0.1505	0.0424	0.1739	0.0320	0.0456	0.1216	0.0169
	$\tau_{max} = 50$	0.3160	0.1924	0.1369	0.0818	0.1572	0.0314	0.0444	0.1278	0.0183
	$\tau_{max} = 100$	0.3317	0.1550	0.1341	0.1288	0.1477	0.0384	0.0623	0.1322	0.0214

Table C.15: **Finite sample properties of DMCA I.** DMCA (*top*) and DMCA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	—	—	—	-0.0050	0.0641	0.0041	-0.0008	0.0410	0.0017
	$\kappa_{max} = 51$	—	—	—	-0.0118	0.0787	0.0063	-0.0064	0.0484	0.0024
	$\kappa_{max} = 101$	—	—	—	-0.0237	0.1017	0.0109	-0.0113	0.0599	0.0037
$T = 1000$	$\kappa_{max} = 21$	—	—	—	-0.0066	0.0466	0.0022	-0.0030	0.0304	0.0009
	$\kappa_{max} = 51$	—	—	—	-0.0151	0.0537	0.0031	-0.0077	0.0350	0.0013
	$\kappa_{max} = 101$	—	—	—	-0.0194	0.0710	0.0054	-0.0074	0.0436	0.0020
$T = 5000$	$\kappa_{max} = 21$	—	—	—	-0.0021	0.0203	0.0004	-0.0021	0.0134	0.0002
	$\kappa_{max} = 51$	—	—	—	-0.0058	0.0228	0.0006	-0.0054	0.0157	0.0003
	$\kappa_{max} = 101$	—	—	—	-0.0066	0.0277	0.0008	-0.0058	0.0190	0.0004
$T = 500$	$\kappa_{max} = 21$	-0.0018	0.0331	0.0011	-0.0031	0.0381	0.0015	-0.0008	0.0397	0.0016
	$\kappa_{max} = 51$	-0.0059	0.0376	0.0015	-0.0067	0.0438	0.0020	-0.0063	0.0467	0.0022
	$\kappa_{max} = 101$	-0.0090	0.0462	0.0022	-0.0105	0.0531	0.0029	-0.0109	0.0578	0.0035
$T = 1000$	$\kappa_{max} = 21$	-0.0030	0.0217	0.0005	-0.0044	0.0273	0.0008	-0.0030	0.0294	0.0009
	$\kappa_{max} = 51$	-0.0072	0.0252	0.0007	-0.0095	0.0300	0.0010	-0.0075	0.0339	0.0012
	$\kappa_{max} = 101$	-0.0071	0.0319	0.0011	-0.0095	0.0377	0.0015	-0.0073	0.0422	0.0018
$T = 5000$	$\kappa_{max} = 21$	-0.0025	0.0102	0.0001	-0.0017	0.0118	0.0001	-0.0021	0.0130	0.0002
	$\kappa_{max} = 51$	-0.0057	0.0115	0.0002	-0.0053	0.0135	0.0002	-0.0054	0.0152	0.0003
	$\kappa_{max} = 101$	-0.0054	0.0139	0.0002	-0.0057	0.0164	0.0003	-0.0058	0.0184	0.0004

Table C.16: **Finite sample properties of DMCA II.** DMCA (*top*) and DMCA_{abs} (*bottom*) estimators for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	—	—	—	-0.0613	0.0718	0.0089	-0.0593	0.0471	0.0057
	$\kappa_{max} = 51$	—	—	—	-0.0508	0.0992	0.0124	-0.0411	0.0583	0.0051
	$\kappa_{max} = 101$	—	—	—	-0.0557	0.1358	0.02156	-0.0344	0.0728	0.0065
$T = 1000$	$\kappa_{max} = 21$	—	—	—	-0.0583	0.0502	0.0059	-0.0550	0.0314	0.0040
	$\kappa_{max} = 51$	—	—	—	-0.0407	0.0613	0.0054	-0.0366	0.0400	0.0029
	$\kappa_{max} = 101$	—	—	—	-0.0367	0.0850	0.0086	-0.0266	0.0516	0.0034
$T = 5000$	$\kappa_{max} = 21$	—	—	—	-0.0569	0.0221	0.0037	-0.0563	0.0141	0.0034
	$\kappa_{max} = 51$	—	—	—	-0.0374	0.0258	0.0021	-0.0349	0.0168	0.0015
	$\kappa_{max} = 101$	—	—	—	-0.0255	0.0313	0.0016	-0.0233	0.0210	0.0010
$T = 500$	$\kappa_{max} = 21$	-0.0574	0.0371	0.0047	-0.0586	0.0432	0.0053	-0.0592	0.0457	0.0056
	$\kappa_{max} = 51$	-0.0395	0.0421	0.0033	-0.0406	0.0530	0.0045	-0.0407	0.0564	0.0048
	$\kappa_{max} = 101$	-0.0310	0.0547	0.0040	-0.0320	0.0654	0.0053	-0.0335	0.0699	0.0060
$T = 1000$	$\kappa_{max} = 21$	-0.0553	0.0241	0.0036	-0.0562	0.0289	0.0040	-0.0550	0.0304	0.0040
	$\kappa_{max} = 51$	-0.0355	0.0292	0.0021	-0.0360	0.0346	0.0025	-0.0366	0.0388	0.0028
	$\kappa_{max} = 101$	-0.0245	0.0372	0.0020	-0.0265	0.0452	0.0027	-0.0264	0.0499	0.0032
$T = 5000$	$\kappa_{max} = 21$	-0.0564	0.0111	0.0033	-0.0564	0.0132	0.0034	-0.0562	0.0137	0.0033
	$\kappa_{max} = 51$	-0.0355	0.0134	0.0014	-0.0358	0.0151	0.0015	-0.0348	0.0163	0.0015
	$\kappa_{max} = 101$	-0.0240	0.0164	0.0008	-0.0239	0.0188	0.0009	-0.0233	0.0204	0.0010

Table C.17: **Finite sample properties of DMCA III.** DMCA (*top*) and DMCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	—	—	—	-0.0274	0.0741	0.0062	-0.0278	0.0453	0.0028
	$\kappa_{max} = 51$	—	—	—	-0.0364	0.0908	0.0106	-0.0316	0.0514	0.0036
	$\kappa_{max} = 101$	—	—	—	-0.0524	0.1349	0.0209	-0.0330	0.0680	0.0057
$T = 1000$	$\kappa_{max} = 21$	—	—	—	-0.0260	0.0487	0.0030	-0.0247	0.0325	0.0017
	$\kappa_{max} = 51$	—	—	—	-0.0321	0.0621	0.0049	-0.0269	0.0381	0.0022
	$\kappa_{max} = 101$	—	—	—	-0.0375	0.0875	0.0091	-0.0259	0.0462	0.0028
$T = 5000$	$\kappa_{max} = 21$	—	—	—	-0.0246	0.0225	0.0011	-0.0238	0.0138	0.0008
	$\kappa_{max} = 51$	—	—	—	-0.0274	0.0277	0.0015	-0.0263	0.0169	0.0010
	$\kappa_{max} = 101$	—	—	—	-0.0256	0.0352	0.0019	-0.0236	0.0207	0.0010
$T = 500$	$\kappa_{max} = 21$	0.0125	0.0339	0.0013	0.0065	0.0389	0.0016	-0.0106	0.0424	0.0019
	$\kappa_{max} = 51$	0.0063	0.0372	0.0014	0.0010	0.0448	0.0020	-0.0139	0.0472	0.0024
	$\kappa_{max} = 101$	0.0013	0.0485	0.0024	-0.0040	0.0550	0.0030	-0.0159	0.0611	0.0040
$T = 1000$	$\kappa_{max} = 21$	0.0152	0.0233	0.0008	0.0065	0.0259	0.0007	-0.0080	0.0305	0.0010
	$\kappa_{max} = 51$	0.0076	0.0274	0.0008	0.0019	0.0305	0.0009	-0.0101	0.0350	0.0013
	$\kappa_{max} = 101$	0.0030	0.0347	0.0012	-0.0008	0.0374	0.0014	-0.0105	0.0424	0.0019
$T = 5000$	$\kappa_{max} = 21$	0.0148	0.0106	0.0003	0.0073	0.0122	0.0002	-0.0074	0.0131	0.0002
	$\kappa_{max} = 51$	0.0089	0.0120	0.0002	0.0026	0.0140	0.0002	-0.0099	0.0156	0.0003
	$\kappa_{max} = 101$	0.0059	0.0143	0.0002	0.0011	0.0168	0.0003	-0.0089	0.0188	0.0004

Table C.18: **Finite sample properties of DMCA IV.** DMCA (*top*) and DMCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	—	—	—	0.1493	0.0725	0.0275	0.1503	0.0478	0.0249
	$\kappa_{max} = 51$	—	—	—	0.0797	0.0967	0.0157	0.0891	0.0521	0.0107
	$\kappa_{max} = 101$	—	—	—	0.0187	0.1344	0.0184	0.0447	0.0699	0.0069
$T = 1000$	$\kappa_{max} = 21$	—	—	—	0.1481	0.0485	0.0243	0.1493	0.0324	0.0233
	$\kappa_{max} = 51$	—	—	—	0.0857	0.0657	0.0117	0.0903	0.0366	0.0095
	$\kappa_{max} = 101$	—	—	—	0.0433	0.0875	0.0095	0.0487	0.0450	0.0044
$T = 5000$	$\kappa_{max} = 21$	—	—	—	0.1502	0.0221	0.0230	0.1521	0.0144	0.0233
	$\kappa_{max} = 51$	—	—	—	0.0927	0.0269	0.0093	0.0933	0.0167	0.0090
	$\kappa_{max} = 101$	—	—	—	0.0548	0.0322	0.0040	0.0546	0.0199	0.0034
$T = 500$	$\kappa_{max} = 21$	0.1632	0.0347	0.0278	0.1623	0.0417	0.0281	0.1549	0.0458	0.0261
	$\kappa_{max} = 51$	0.1172	0.0383	0.0152	0.1114	0.0480	0.0147	0.1003	0.0492	0.0125
	$\kappa_{max} = 101$	0.0832	0.0487	0.0093	0.0726	0.0559	0.0084	0.0604	0.0651	0.0079
$T = 1000$	$\kappa_{max} = 21$	0.1636	0.0253	0.0274	0.1603	0.0283	0.0265	0.1538	0.0312	0.0246
	$\kappa_{max} = 51$	0.1186	0.0287	0.0149	0.1119	0.0337	0.0137	0.1009	0.0348	0.0114
	$\kappa_{max} = 101$	0.0843	0.0337	0.0082	0.0790	0.0405	0.0079	0.0632	0.0419	0.0058
$T = 5000$	$\kappa_{max} = 21$	0.1648	0.0112	0.0273	0.1603	0.0127	0.0259	0.1564	0.0138	0.0247
	$\kappa_{max} = 51$	0.1198	0.0124	0.0145	0.1139	0.0150	0.0132	0.1034	0.0158	0.0109
	$\kappa_{max} = 101$	0.0867	0.0149	0.0077	0.0817	0.0169	0.0070	0.0680	0.0187	0.0050

Table C.19: **Finite sample properties of DMCA V.** DMCA (*top*) and DMCA_{abs} (*bottom*) estimators for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	—	—	—	0.3160	0.0759	0.1056	0.3225	0.0472	0.1062
	$\kappa_{max} = 51$	—	—	—	0.2717	0.0939	0.0826	0.2761	0.0557	0.0793
	$\kappa_{max} = 101$	—	—	—	0.1920	0.1324	0.0544	0.2087	0.0700	0.0485
$T = 1000$	$\kappa_{max} = 21$	—	—	—	0.3288	0.0534	0.1078	0.3255	0.0339	0.1071
	$\kappa_{max} = 51$	—	—	—	0.2771	0.0618	0.0806	0.2837	0.0388	0.0820
	$\kappa_{max} = 101$	—	—	—	0.2074	0.0799	0.0494	0.2171	0.0481	0.0493
$T = 5000$	$\kappa_{max} = 21$	—	—	—	0.3268	0.0229	0.1074	0.3269	0.0151	0.1071
	$\kappa_{max} = 51$	—	—	—	0.2843	0.0264	0.0815	0.2853	0.0172	0.0817
	$\kappa_{max} = 101$	—	—	—	0.2181	0.0329	0.0487	0.2204	0.0204	0.0490
$T = 500$	$\kappa_{max} = 21$	0.3130	0.0362	0.0999	0.3131	0.0426	0.0999	0.3184	0.0447	0.1034
	$\kappa_{max} = 51$	0.2667	0.0409	0.0728	0.2700	0.0507	0.0755	0.2718	0.0530	0.0767
	$\kappa_{max} = 101$	0.2101	0.0526	0.0469	0.2116	0.0628	0.0487	0.2099	0.0661	0.0484
$T = 1000$	$\kappa_{max} = 21$	0.3155	0.0236	0.1002	0.3177	0.0304	0.1018	0.3212	0.0321	0.1042
	$\kappa_{max} = 51$	0.2728	0.0303	0.0753	0.2730	0.0344	0.0757	0.2789	0.0372	0.0792
	$\kappa_{max} = 101$	0.2189	0.0366	0.0493	0.2164	0.0426	0.0486	0.2174	0.0448	0.0493
$T = 5000$	$\kappa_{max} = 21$	0.3157	0.0114	0.0998	0.3187	0.0130	0.1017	0.3225	0.0143	0.1042
	$\kappa_{max} = 51$	0.2719	0.0133	0.0741	0.2746	0.0150	0.0756	0.2802	0.0165	0.0788
	$\kappa_{max} = 101$	0.2197	0.0157	0.0485	0.2186	0.0180	0.0481	0.2200	0.0195	0.0488

Table C.20: **Finite sample properties of DMCA VI.** DMCA_{abs} estimator for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	0.0649	0.0345	0.0054	0.0604	0.0365	0.0049	0.0499	0.0379	0.0039
	$\kappa_{max} = 51$	0.0864	0.0399	0.0091	0.0840	0.0419	0.0088	0.0715	0.0436	0.0070
	$\kappa_{max} = 101$	0.1012	0.0536	0.0131	0.0983	0.0521	0.0124	0.0881	0.0544	0.0107
$T = 1000$	$\kappa_{max} = 21$	0.0640	0.0241	0.0047	0.0596	0.0248	0.0042	0.0515	0.0282	0.0034
	$\kappa_{max} = 51$	0.0882	0.0295	0.0086	0.0828	0.0288	0.0077	0.0751	0.0310	0.0066
	$\kappa_{max} = 101$	0.1055	0.0362	0.0124	0.0998	0.0348	0.0112	0.0925	0.0382	0.0100
$T = 5000$	$\kappa_{max} = 21$	0.0645	0.0106	0.0043	0.0610	0.0108	0.0038	0.0509	0.0123	0.0027
	$\kappa_{max} = 51$	0.0892	0.0128	0.0081	0.0850	0.0123	0.0074	0.0756	0.0138	0.0059
	$\kappa_{max} = 101$	0.1069	0.0154	0.0117	0.1022	0.0157	0.0107	0.0949	0.0174	0.0093

Table C.21: **Finite sample properties of DMCA VII.** DMCA estimator of H_ρ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
		bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$\kappa_{max} = 21$	0.2652	0.4064	0.2355	0.0723	0.1976	0.0443	0.0973	0.0834	0.0164
	$\kappa_{max} = 51$	0.3622	0.3062	0.2249	0.0360	0.2404	0.0591	0.0442	0.1759	0.0329
	$\kappa_{max} = 101$	0.4117	0.2580	0.2361	0.0920	0.2202	0.0570	0.0293	0.1998	0.0408
$T = 1000$	$\kappa_{max} = 21$	0.1942	0.4033	0.2004	0.0891	0.1229	0.0230	0.1062	0.0540	0.0142
	$\kappa_{max} = 51$	0.2893	0.2994	0.1733	0.0355	0.2106	0.0456	0.0721	0.1022	0.0156
	$\kappa_{max} = 101$	0.3453	0.2706	0.1925	0.0440	0.2148	0.0481	0.0306	0.1749	0.0315
$T = 5000$	$\kappa_{max} = 21$	0.0640	0.2816	0.0834	0.1056	0.0475	0.0134	0.1096	0.0223	0.0125
	$\kappa_{max} = 51$	0.1097	0.2658	0.0826	0.0821	0.0799	0.0131	0.0917	0.0369	0.0098
	$\kappa_{max} = 101$	0.1806	0.2358	0.0882	0.0494	0.1402	0.0221	0.0732	0.0644	0.0095

Table C.22: **Finite sample properties of APE I.** APE estimator for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1\text{T}$	11	-0.0267	0.1360	0.0192	-0.0394	0.1314	0.0188	-0.0305	0.0931	0.0096
	$m = 0.1\text{T}$	21	-0.0398	0.1697	0.0304	-0.0449	0.1316	0.0193	-0.0317	0.0847	0.0082
	$m = 0.1\text{T}$	51	-0.0841	0.1980	0.0463	-0.0585	0.0963	0.0127	-0.0439	0.0570	0.0052
	$m = 0.4\text{T}$	11	-0.0221	0.0742	0.0060	-0.0194	0.0687	0.0051	-0.0179	0.0515	0.0030
	$m = 0.4\text{T}$	21	-0.0281	0.0992	0.0106	-0.0206	0.0730	0.0058	-0.0179	0.0504	0.0029
	$m = 0.4\text{T}$	51	-0.0386	0.1440	0.0222	-0.0210	0.0718	0.0056	-0.0182	0.0471	0.0026
$T = 1000$	$m = 0.1\text{T}$	11	-0.0251	0.1010	0.0108	-0.0181	0.0965	0.0096	-0.0144	0.0681	0.0048
	$m = 0.1\text{T}$	21	-0.0306	0.1316	0.0183	-0.0207	0.1003	0.0105	-0.0149	0.0658	0.0046
	$m = 0.1\text{T}$	51	-0.0528	0.1811	0.0356	-0.0236	0.0903	0.0087	-0.0182	0.0579	0.0037
	$m = 0.4\text{T}$	11	-0.0136	0.0512	0.0028	-0.0161	0.0491	0.0027	-0.0114	0.0367	0.0015
	$m = 0.4\text{T}$	21	-0.0153	0.0688	0.0050	-0.0169	0.0516	0.0029	-0.0116	0.0365	0.0015
	$m = 0.4\text{T}$	51	-0.0220	0.1010	0.0107	-0.0176	0.0517	0.0029	-0.0119	0.0357	0.0014
$T = 5000$	$m = 0.1\text{T}$	11	-0.0044	0.0450	0.0020	-0.0029	0.0406	0.0017	-0.0047	0.0314	0.0010
	$m = 0.1\text{T}$	21	-0.0077	0.0592	0.0036	-0.0030	0.0430	0.0019	-0.0046	0.0312	0.0010
	$m = 0.1\text{T}$	51	-0.0150	0.0901	0.0083	-0.0032	0.0433	0.0019	-0.0045	0.0305	0.0010
	$m = 0.4\text{T}$	11	-0.0090	0.0226	0.0006	-0.0099	0.0211	0.0005	-0.0101	0.0155	0.0003
	$m = 0.4\text{T}$	21	-0.0096	0.0298	0.0010	-0.0102	0.0224	0.0006	-0.0101	0.0155	0.0003
	$m = 0.4\text{T}$	51	-0.0107	0.0459	0.0022	-0.0103	0.0231	0.0006	-0.0101	0.0155	0.0003

Table C.23: **Finite sample properties of APE II.** APE estimator for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	-0.1002	0.0826	0.0169	-0.1125	0.0875	0.0203	-0.0967	0.0632	0.0133
	$m = 0.1$T	21	-0.1085	0.1067	0.0232	-0.1208	0.0968	0.0240	-0.0991	0.0618	0.0136
	$m = 0.1$T	51	-0.1910	0.1788	0.0685	-0.1740	0.1199	0.0447	-0.1418	0.0574	0.0234
	$m = 0.4$T	11	-0.0750	0.0454	0.0077	-0.0821	0.0468	0.0089	-0.0753	0.0333	0.0068
	$m = 0.4$T	21	-0.0747	0.0586	0.0090	-0.0833	0.0509	0.0095	-0.0745	0.0336	0.0067
	$m = 0.4$T	51	-0.0786	0.0833	0.0131	-0.0849	0.0544	0.0102	-0.0746	0.0331	0.0067
$T = 1000$	$m = 0.1$T	11	-0.0772	0.0608	0.0097	-0.0839	0.0604	0.0107	-0.0778	0.0429	0.0079
	$m = 0.1$T	21	-0.0795	0.0792	0.0126	-0.0876	0.0690	0.0124	-0.0774	0.0429	0.0078
	$m = 0.1$T	51	-0.0930	0.1139	0.0216	-0.0971	0.0786	0.0156	-0.0821	0.0419	0.0085
	$m = 0.4$T	11	-0.0626	0.0341	0.0051	-0.0659	0.0324	0.0054	-0.0650	0.0251	0.0049
	$m = 0.4$T	21	-0.0625	0.0441	0.0059	-0.0665	0.0355	0.0057	-0.0643	0.0254	0.0048
	$m = 0.4$T	51	-0.0635	0.0629	0.0080	-0.0666	0.0379	0.0059	-0.0637	0.0254	0.0047
$T = 5000$	$m = 0.1$T	11	-0.0439	0.0285	0.0027	-0.0490	0.0272	0.0031	-0.0466	0.0206	0.0026
	$m = 0.1$T	21	-0.0433	0.0373	0.0033	-0.0495	0.0298	0.0033	-0.0460	0.0208	0.0025
	$m = 0.1$T	51	-0.0442	0.0546	0.0049	-0.0495	0.0318	0.0035	-0.0454	0.0209	0.0025
	$m = 0.4$T	11	-0.0439	0.0185	0.0023	-0.0456	0.0173	0.0024	-0.0455	0.0133	0.0023
	$m = 0.4$T	21	-0.0430	0.0239	0.0024	-0.0457	0.0192	0.0025	-0.0451	0.0136	0.0022
	$m = 0.4$T	51	-0.0418	0.0340	0.0029	-0.0455	0.0206	0.0025	-0.0446	0.0138	0.0022

Table C.24: **Finite sample properties of APE III.** APE estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	-0.0428	0.1172	0.0156	-0.0498	0.1130	0.0152	-0.0481	0.0864	0.0098
	$m = 0.1$T	21	-0.0580	0.1500	0.0259	-0.0678	0.1202	0.0190	-0.0588	0.0811	0.0100
	$m = 0.1$T	51	-0.1205	0.1880	0.0499	-0.1294	0.1130	0.0295	-0.1103	0.0599	0.0158
	$m = 0.4$T	11	-0.0052	0.0593	0.0035	-0.0088	0.0519	0.0028	-0.0064	0.0408	0.0017
	$m = 0.4$T	21	-0.0091	0.0776	0.0061	-0.0134	0.0556	0.0033	-0.0094	0.0408	0.0018
	$m = 0.4$T	51	-0.0225	0.1118	0.0130	-0.0230	0.0577	0.0039	-0.0171	0.0396	0.0019
$T = 1000$	$m = 0.1$T	11	-0.0245	0.0857	0.0079	-0.0293	0.0782	0.0070	-0.0326	0.0594	0.0046
	$m = 0.1$T	21	-0.0323	0.1147	0.0142	-0.0376	0.0832	0.0083	-0.0384	0.0582	0.0049
	$m = 0.1$T	51	-0.0607	0.1598	0.0292	-0.0567	0.0823	0.0100	-0.0562	0.0530	0.0060
	$m = 0.4$T	11	0.0022	0.0412	0.0017	0.0029	0.0370	0.0014	0.0010	0.0287	0.0008
	$m = 0.4$T	21	-0.0011	0.0554	0.0031	0.0001	0.0394	0.0016	-0.0010	0.0287	0.0008
	$m = 0.4$T	51	-0.0107	0.0864	0.0076	-0.0048	0.0408	0.0017	-0.0055	0.0283	0.0008
$T = 5000$	$m = 0.1$T	11	-0.0073	0.0383	0.0015	-0.0101	0.0345	0.0013	-0.0094	0.0265	0.0008
	$m = 0.1$T	21	-0.0099	0.0521	0.0028	-0.0126	0.0370	0.0015	-0.0111	0.0265	0.0008
	$m = 0.1$T	51	-0.0178	0.0820	0.0070	-0.0172	0.0385	0.0018	-0.0150	0.0263	0.0009
	$m = 0.4$T	11	0.0107	0.0192	0.0005	0.0092	0.0177	0.0004	0.0098	0.0129	0.0003
	$m = 0.4$T	21	0.0098	0.0260	0.0008	0.0083	0.0188	0.0004	0.0091	0.0129	0.0002
	$m = 0.4$T	51	0.0080	0.0393	0.0016	0.0065	0.0194	0.0004	0.0075	0.0129	0.0002

Table C.25: **Finite sample properties of APE IV.** APE estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	-0.0072	0.1069	0.0115	-0.0241	0.1056	0.0117	-0.0142	0.0781	0.0063
	$m = 0.1$T	21	-0.0198	0.1334	0.0182	-0.0379	0.1098	0.0135	-0.0240	0.0742	0.0061
	$m = 0.1$T	51	-0.0967	0.1839	0.0432	-0.0865	0.1044	0.0184	-0.0668	0.0555	0.0075
	$m = 0.4$T	11	0.1114	0.0419	0.0142	0.1078	0.0386	0.0131	0.1113	0.0282	0.0132
	$m = 0.4$T	21	0.1071	0.0559	0.0146	0.1039	0.0421	0.0126	0.1090	0.0282	0.0127
	$m = 0.4$T	51	0.0957	0.0816	0.0158	0.0967	0.0451	0.0114	0.1034	0.0281	0.0115
$T = 1000$	$m = 0.1$T	11	0.0080	0.0813	0.0067	0.0035	0.0745	0.0056	0.0030	0.0547	0.0030
	$m = 0.1$T	21	-0.0011	0.1071	0.0115	-0.0036	0.0795	0.0063	-0.0016	0.0539	0.0029
	$m = 0.1$T	51	-0.0309	0.1501	0.0235	-0.0211	0.0823	0.0072	-0.0153	0.0495	0.0027
	$m = 0.4$T	11	0.1166	0.0284	0.0144	0.1141	0.0279	0.0138	0.1164	0.0199	0.0140
	$m = 0.4$T	21	0.1151	0.0381	0.0147	0.1119	0.0299	0.0134	0.1150	0.0199	0.0136
	$m = 0.4$T	51	0.1089	0.0587	0.0153	0.1080	0.0313	0.0126	0.1119	0.0199	0.0129
$T = 5000$	$m = 0.1$T	11	0.0271	0.0350	0.0020	0.0288	0.0315	0.0018	0.0247	0.0244	0.0012
	$m = 0.1$T	21	0.0257	0.0455	0.0027	0.0268	0.0338	0.0019	0.0231	0.0244	0.0011
	$m = 0.1$T	51	0.0203	0.0704	0.0054	0.0227	0.0349	0.0017	0.0194	0.0239	0.0009
	$m = 0.4$T	11	0.1227	0.0133	0.0152	0.1229	0.0117	0.0153	0.1224	0.0089	0.0151
	$m = 0.4$T	21	0.1226	0.0179	0.0154	0.1223	0.0124	0.0151	0.1219	0.0089	0.0149
	$m = 0.4$T	51	0.1213	0.0266	0.01543	0.1211	0.0128	0.0148	0.1207	0.0090	0.0146

Table C.26: **Finite sample properties of APE V.** APE estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$, $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.0896	0.0831	0.0149	0.0834	0.0852	0.0142	0.0883	0.0559	0.0109
	$m = 0.1$T	21	0.0725	0.1087	0.0171	0.0690	0.0941	0.0136	0.0792	0.0551	0.0093
	$m = 0.1$T	51	-0.0336	0.1803	0.0336	0.0077	0.1032	0.0107	0.0249	0.0482	0.0029
	$m = 0.4$T	11	0.1945	0.0280	0.0386	0.1890	0.0275	0.0365	0.1914	0.0200	0.0371
	$m = 0.4$T	21	0.1919	0.0366	0.0381	0.1860	0.0309	0.0355	0.1902	0.0202	0.0366
	$m = 0.4$T	51	0.1844	0.0561	0.0372	0.1824	0.0347	0.0345	0.1880	0.0202	0.0357
$T = 1000$	$m = 0.1$T	11	0.1081	0.0573	0.0150	0.1017	0.0563	0.0135	0.1042	0.0403	0.0125
	$m = 0.1$T	21	0.1013	0.0745	0.0158	0.0947	0.0617	0.0128	0.1002	0.0398	0.0116
	$m = 0.1$T	51	0.0728	0.1172	0.0190	0.0771	0.0695	0.0108	0.0871	0.0383	0.0091
	$m = 0.4$T	11	0.1976	0.0200	0.0395	0.1962	0.0178	0.0388	0.1966	0.0133	0.0388
	$m = 0.4$T	21	0.1961	0.0268	0.0392	0.1943	0.0198	0.0382	0.1957	0.0134	0.0385
	$m = 0.4$T	51	0.1923	0.0403	0.0386	0.1919	0.0223	0.0373	0.1942	0.0134	0.0379
$T = 5000$	$m = 0.1$T	11	0.1217	0.0254	0.0154	0.1208	0.0240	0.0152	0.1197	0.0185	0.0147
	$m = 0.1$T	21	0.1206	0.0340	0.0157	0.1193	0.0258	0.0149	0.1185	0.0185	0.0144
	$m = 0.1$T	51	0.1169	0.0527	0.0165	0.1160	0.0272	0.0142	0.1158	0.0184	0.0137
	$m = 0.4$T	11	0.2038	0.0086	0.0416	0.2028	0.0081	0.0412	0.2030	0.0062	0.0412
	$m = 0.4$T	21	0.2036	0.0116	0.0416	0.2021	0.0087	0.0409	0.2025	0.0063	0.0411
	$m = 0.4$T	51	0.2026	0.0182	0.0414	0.2010	0.0091	0.0405	0.2016	0.0063	0.0407

Table C.27: **Finite sample properties of APE VI.** APE estimator for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.0599	0.0965	0.0129	0.0405	0.1027	0.0122	0.0070	0.1108	0.0123
	$m = 0.1$T	21	0.0542	0.1230	0.0181	0.0188	0.1281	0.0168	-0.0274	0.1390	0.0201
	$m = 0.1$T	51	-0.0142	0.1791	0.0323	-0.0871	0.2258	0.0586	-0.1459	0.2371	0.0775
	$m = 0.4$T	11	0.0674	0.0512	0.0072	0.0380	0.0585	0.0049	0.0102	0.0586	0.0035
	$m = 0.4$T	21	0.0672	0.0679	0.0091	0.0198	0.0764	0.0062	-0.0089	0.0704	0.0050
	$m = 0.4$T	51	0.0662	0.0981	0.0140	-0.0168	0.1094	0.0122	-0.0325	0.0860	0.0085
$T = 1000$	$m = 0.1$T	11	0.0833	0.0671	0.0114	0.0625	0.0738	0.0094	0.0328	0.0825	0.0079
	$m = 0.1$T	21	0.0829	0.0875	0.0145	0.0450	0.0976	0.0116	0.0067	0.1033	0.0107
	$m = 0.1$T	51	0.0749	0.1217	0.0204	-0.0031	0.1403	0.0197	-0.0442	0.1390	0.0213
	$m = 0.4$T	11	0.0765	0.0386	0.0073	0.0501	0.0438	0.0044	0.0192	0.0446	0.0024
	$m = 0.4$T	21	0.0766	0.0504	0.0084	0.0335	0.0570	0.0044	0.0021	0.0523	0.0027
	$m = 0.4$T	51	0.0751	0.0770	0.0116	0.0066	0.0768	0.0059	-0.0176	0.0630	0.0043
$T = 5000$	$m = 0.1$T	11	0.1136	0.0325	0.0140	0.0953	0.0343	0.0103	0.0668	0.0372	0.0059
	$m = 0.1$T	21	0.1142	0.0424	0.0148	0.0831	0.0468	0.0091	0.0463	0.0471	0.0044
	$m = 0.1$T	51	0.1159	0.0604	0.0171	0.0573	0.0709	0.0083	0.0192	0.0621	0.0042
	$m = 0.4$T	11	0.0959	0.0208	0.0096	0.0682	0.0224	0.0052	0.0359	0.0218	0.0018
	$m = 0.4$T	21	0.0966	0.0273	0.0101	0.0515	0.0299	0.0036	0.0177	0.0263	0.0010
	$m = 0.4$T	51	0.0977	0.0390	0.0111	0.0260	0.0412	0.0024	-0.0003	0.0313	0.0010

Table C.28: **Finite sample properties of APE VII.** APE estimator of $\rho_H = H_{xy} - \frac{H_x+H_y}{2}$ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.2013	0.0834	0.0475	0.1818	0.0904	0.0412	0.1502	0.0887	0.0304
	$m = 0.1$T	21	0.1965	0.1095	0.0506	0.1575	0.1273	0.0410	0.1154	0.1211	0.0280
	$m = 0.1$T	51	0.1643	0.1733	0.0570	0.0790	0.2101	0.0504	0.0296	0.2433	0.0601
	$m = 0.4$T	11	0.2013	0.0449	0.0426	0.1761	0.0458	0.0331	0.1428	0.0451	0.0224
	$m = 0.4$T	21	0.2007	0.0632	0.0443	0.1582	0.0658	0.0294	0.1234	0.0579	0.0186
	$m = 0.4$T	51	0.1965	0.0958	0.0478	0.1253	0.0992	0.0255	0.1007	0.0762	0.0159
$T = 1000$	$m = 0.1$T	11	0.2048	0.0550	0.0450	0.1817	0.0602	0.0366	0.1506	0.0620	0.0265
	$m = 0.1$T	21	0.2035	0.0763	0.0472	0.1638	0.0845	0.0340	0.1240	0.0838	0.0224
	$m = 0.1$T	51	0.2006	0.1132	0.0531	0.1236	0.1289	0.0319	0.0793	0.1250	0.0219
	$m = 0.4$T	11	0.2007	0.0324	0.0413	0.1740	0.0355	0.0315	0.1417	0.0333	0.0212
	$m = 0.4$T	21	0.2014	0.0441	0.0425	0.1569	0.0496	0.0271	0.1224	0.0426	0.0168
	$m = 0.4$T	51	0.2020	0.0689	0.0455	0.1270	0.0730	0.0215	0.1011	0.0561	0.0134
$T = 5000$	$m = 0.1$T	11	0.2008	0.0267	0.0410	0.1836	0.0280	0.0345	0.1548	0.0299	0.0249
	$m = 0.1$T	21	0.2014	0.0364	0.0419	0.1704	0.0407	0.0307	0.1330	0.0418	0.0194
	$m = 0.1$T	51	0.2035	0.0555	0.0445	0.1444	0.0640	0.0250	0.1042	0.0586	0.0143
	$m = 0.4$T	11	0.2001	0.0161	0.0403	0.1747	0.0176	0.0308	0.1418	0.0168	0.0204
	$m = 0.4$T	21	0.1997	0.0231	0.0404	0.1586	0.0245	0.0257	0.1238	0.0213	0.0158
	$m = 0.4$T	51	0.1992	0.0366	0.0410	0.1336	0.0356	0.0191	0.1062	0.0272	0.0120

Table C.29: **Finite sample properties of XPE I.** XPE for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	-0.0217	0.1305	0.0175	-0.0411	0.1318	0.0191	-0.0292	0.0899	0.0089
	$m = 0.1T$	21	-0.0376	0.1658	0.0289	-0.0489	0.1295	0.0192	-0.0323	0.0800	0.0074
	$m = 0.1T$	51	-0.0776	0.1763	0.0371	-0.0611	0.0734	0.0091	-0.0535	0.0472	0.0051
	$m = 0.4T$	11	-0.0115	0.0659	0.0045	-0.0200	0.0619	0.0042	-0.0169	0.0443	0.0023
	$m = 0.4T$	21	-0.0170	0.0860	0.0077	-0.0222	0.0646	0.0047	-0.0173	0.0431	0.0022
	$m = 0.4T$	51	-0.0307	0.1223	0.01589	-0.0255	0.0578	0.0040	-0.0215	0.0400	0.0021
$T = 1000$	$m = 0.1T$	11	-0.0157	0.0937	0.0090	-0.0176	0.0938	0.0091	-0.0139	0.0608	0.0039
	$m = 0.1T$	21	-0.0237	0.1189	0.0147	-0.0220	0.0954	0.0096	-0.0150	0.0579	0.0036
	$m = 0.1T$	51	-0.0541	0.1593	0.0283	-0.0310	0.0781	0.0071	-0.0257	0.0482	0.0030
	$m = 0.4T$	11	-0.0113	0.0450	0.0022	-0.0125	0.0436	0.0021	-0.0112	0.0301	0.0010
	$m = 0.4T$	21	-0.0135	0.0605	0.0038	-0.0146	0.0463	0.0024	-0.0112	0.0295	0.0010
	$m = 0.4T$	51	-0.0167	0.0931	0.0089	-0.0168	0.0443	0.0022	-0.0132	0.0284	0.0010
$T = 5000$	$m = 0.1T$	11	-0.0037	0.0397	0.0016	-0.0040	0.0399	0.0016	-0.0053	0.0263	0.0007
	$m = 0.1T$	21	-0.0041	0.0524	0.0028	-0.0050	0.0416	0.0018	-0.0055	0.0259	0.0007
	$m = 0.1T$	51	-0.0106	0.0812	0.0067	-0.0058	0.0393	0.0016	-0.0071	0.0251	0.0007
	$m = 0.4T$	11	-0.0068	0.0197	0.0004	-0.0082	0.0200	0.0005	-0.0073	0.0134	0.0002
	$m = 0.4T$	21	-0.0071	0.0269	0.0008	-0.0083	0.0206	0.0005	-0.0073	0.0133	0.0002
	$m = 0.4T$	51	-0.0085	0.0422	0.0019	-0.0085	0.0202	0.0005	-0.0076	0.0132	0.0002

Table C.30: **Finite sample properties of XPE II.** XPE for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	-0.0243	0.1342	0.0186	-0.0573	0.1444	0.0241	-0.0435	0.1018	0.0122
	$m = 0.1T$	21	-0.0456	0.1821	0.0352	-0.0833	0.1583	0.0320	-0.0600	0.0989	0.0134
	$m = 0.1T$	51	-0.1685	0.2078	0.0716	-0.1809	0.1069	0.0441	-0.1652	0.0669	0.0318
	$m = 0.4T$	11	-0.0238	0.0653	0.0048	-0.0399	0.0667	0.0060	-0.0365	0.0472	0.0036
	$m = 0.4T$	21	-0.0174	0.0871	0.0079	-0.0424	0.0775	0.0078	-0.0345	0.0488	0.0036
	$m = 0.4T$	51	-0.0237	0.1445	0.0215	-0.0480	0.0867	0.0098	-0.0392	0.0519	0.0042
$T = 1000$	$m = 0.1T$	11	-0.0062	0.0946	0.0090	-0.0266	0.0930	0.0093	-0.0214	0.0689	0.0052
	$m = 0.1T$	21	-0.0053	0.1323	0.0175	-0.0356	0.1070	0.0127	-0.0234	0.0700	0.0054
	$m = 0.1T$	51	-0.0370	0.1987	0.0408	-0.0681	0.1050	0.0156	-0.0565	0.0668	0.0077
	$m = 0.4T$	11	-0.0255	0.0436	0.0026	-0.0333	0.0463	0.0033	-0.0304	0.0305	0.0019
	$m = 0.4T$	21	-0.0207	0.0601	0.0040	-0.0336	0.0522	0.0039	-0.0277	0.0317	0.0018
	$m = 0.4T$	51	-0.0140	0.0994	0.0101	-0.0317	0.0574	0.0043	-0.0239	0.0347	0.0018
$T = 5000$	$m = 0.1T$	11	-0.0014	0.0391	0.0015	-0.0048	0.0403	0.0016	-0.0040	0.0283	0.0008
	$m = 0.1T$	21	0.0030	0.0548	0.0030	-0.0048	0.0453	0.0021	-0.0016	0.0291	0.0009
	$m = 0.1T$	51	0.0098	0.0921	0.0086	-0.0018	0.0515	0.0027	-0.0028	0.0318	0.0010
	$m = 0.4T$	11	-0.0240	0.0189	0.0009	-0.0262	0.0200	0.0011	-0.0249	0.0131	0.0008
	$m = 0.4T$	21	-0.0210	0.0264	0.0011	-0.0258	0.0216	0.0011	-0.0236	0.0131	0.0007
	$m = 0.4T$	51	-0.0166	0.0417	0.0020	-0.0231	0.0241	0.0011	-0.0200	0.0139	0.0006

Table C.31: **Finite sample properties of XPE III.**XPE for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	-0.0255	0.1330	0.0183	-0.0580	0.1308	0.0205	-0.0448	0.0910	0.0103
	$m = 0.1T$	21	-0.0458	0.1660	0.0297	-0.0917	0.1387	0.0276	-0.0670	0.0824	0.0113
	$m = 0.1T$	51	-0.1224	0.1827	0.0484	-0.1474	0.0984	0.0314	-0.1289	0.0501	0.0191
	$m = 0.4T$	11	0.0086	0.0632	0.0041	-0.0060	0.0651	0.0043	-0.0026	0.0431	0.0019
	$m = 0.4T$	21	0.0045	0.0839	0.0071	-0.0181	0.0721	0.0055	-0.0094	0.0426	0.0019
	$m = 0.4T$	51	-0.0169	0.1251	0.0159	-0.0371	0.0741	0.0069	-0.0279	0.0409	0.0024
$T = 1000$	$m = 0.1T$	11	-0.0142	-0.0941	0.0091	-0.0285	0.0958	0.0100	-0.0290	0.0633	0.0049
	$m = 0.1T$	21	-0.0272	0.1194	0.0150	-0.0476	0.1037	0.0130	-0.0411	0.0613	0.0055
	$m = 0.1T$	51	-0.0687	0.1699	0.0336	-0.0852	0.0995	0.0172	-0.0765	0.0531	0.0087
	$m = 0.4T$	11	0.0105	0.0463	0.0023	0.0043	0.0433	0.0019	0.0033	0.0286	0.0008
	$m = 0.4T$	21	0.0091	0.0630	0.0040	-0.0023	0.0476	0.0023	-0.0001	0.0284	0.0008
	$m = 0.4T$	51	0.0010	0.0948	0.0090	-0.0125	0.0493	0.0026	-0.0100	0.0279	0.0009
$T = 5000$	$m = 0.1T$	11	-0.0026	0.0420	0.0018	-0.0049	0.0392	0.0016	-0.0059	0.0259	0.0007
	$m = 0.1T$	21	-0.0060	0.0558	0.0032	-0.0100	0.0422	0.0019	-0.0088	0.0258	0.0007
	$m = 0.1T$	51	-0.0161	0.0835	0.0072	-0.0198	0.0435	0.0023	-0.0174	0.0256	0.0010
	$m = 0.4T$	11	0.0125	0.0196	0.0005	0.0107	0.0187	0.0005	0.0100	0.0131	0.0003
	$m = 0.4T$	21	0.0116	0.0266	0.0008	0.0089	0.0195	0.0005	0.0091	0.0130	0.0003
	$m = 0.4T$	51	0.0081	0.0420	0.0018	0.0058	0.0206	0.0005	0.0067	0.0131	0.0002

Table C.32: **Finite sample properties of XPE IV.** XPE for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	0.0084	0.1298	0.0169	-0.0147	0.1320	0.0176	-0.0134	0.0925	0.0087
	$m = 0.1T$	21	-0.0192	0.1663	0.0280	-0.0511	0.1404	0.0223	-0.0322	0.0846	0.0082
	$m = 0.1T$	51	-0.1067	0.1889	0.0471	-0.1038	0.0961	0.0200	-0.0878	0.0505	0.0102
	$m = 0.4T$	11	0.1422	0.0635	0.0243	0.1318	0.0631	0.0214	0.1351	0.0461	0.0204
	$m = 0.4T$	21	0.1348	0.0857	0.0255	0.1211	0.0719	0.0198	0.1295	0.0452	0.0188
	$m = 0.4T$	51	0.1139	0.1344	0.0310	0.1039	0.0732	0.0162	0.1133	0.0434	0.0147
$T = 1000$	$m = 0.1T$	11	0.0227	0.0918	0.0089	0.0058	0.0913	0.0084	0.0067	0.0629	0.0040
	$m = 0.1T$	21	0.0102	0.1219	0.0150	-0.0139	0.1002	0.0102	-0.0043	0.0608	0.0037
	$m = 0.1T$	51	-0.0320	0.1745	0.0315	-0.0475	0.0934	0.0110	-0.0360	0.0524	0.0040
	$m = 0.4T$	11	0.1447	0.0446	0.0229	0.1381	0.0429	0.0209	0.1376	0.0300	0.0198
	$m = 0.4T$	21	0.1415	0.0607	0.0237	0.1314	0.0465	0.0194	0.1344	0.0298	0.0189
	$m = 0.4T$	51	0.1344	0.0909	0.0263	0.1216	0.0479	0.0171	0.1255	0.0293	0.0166
$T = 5000$	$m = 0.1T$	11	0.0326	0.0411	0.0028	0.0251	0.0387	0.0021	0.0266	0.0260	0.0014
	$m = 0.1T$	21	0.0317	0.0548	0.0040	0.0198	0.0419	0.0021	0.0234	0.0259	0.0012
	$m = 0.1T$	51	0.0250	0.0847	0.0078	0.0104	0.0434	0.0020	0.0151	0.0258	0.0009
	$m = 0.4T$	11	0.1430	0.0198	0.0208	0.1411	0.0194	0.0203	0.1411	0.0127	0.0201
	$m = 0.4T$	21	0.1426	0.0264	0.0210	0.1396	0.0201	0.0199	0.1403	0.0125	0.0198
	$m = 0.4T$	51	0.1398	0.0420	0.0213	0.1369	0.0209	0.0192	0.1380	0.0126	0.0192

Table C.33: **Finite sample properties of XPE V.** XPE for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	0.1482	0.1354	0.0403	0.1015	0.1387	0.0295	0.1267	0.0933	0.0248
	$m = 0.1T$	21	0.1153	0.1737	0.0435	0.0664	0.1578	0.0293	0.1074	0.0863	0.0190
	$m = 0.1T$	51	-0.0202	0.1985	0.0398	-0.0117	0.1079	0.0118	0.0172	0.0538	0.0032
	$m = 0.4T$	11	0.2960	0.0655	0.0919	0.2847	0.0697	0.0859	0.2910	0.0428	0.0865
	$m = 0.4T$	21	0.2968	0.0874	0.0957	0.2760	0.0778	0.0822	0.2882	0.0425	0.0848
	$m = 0.4T$	51	0.2837	0.1360	0.0990	0.2644	0.0821	0.0766	0.2772	0.0429	0.0787
$T = 1000$	$m = 0.1T$	11	0.1483	0.0921	0.0305	0.1388	0.0910	0.0275	0.1368	0.0645	0.0229
	$m = 0.1T$	21	0.1392	0.1191	0.0336	0.1197	0.1000	0.0243	0.1269	0.0617	0.0199
	$m = 0.1T$	51	0.0799	0.1765	0.0375	0.0829	0.0939	0.0156	0.0925	0.0534	0.0114
	$m = 0.4T$	11	0.2939	0.0425	0.0882	0.2872	0.0438	0.0844	0.2868	0.0293	0.0831
	$m = 0.4T$	21	0.2919	0.0589	0.0887	0.2819	0.0480	0.0818	0.2848	0.0291	0.0819
	$m = 0.4T$	51	0.2853	0.0952	0.0904	0.2763	0.0526	0.0791	0.2804	0.0293	0.0795
$T = 5000$	$m = 0.1T$	11	0.1475	0.0383	0.0232	0.1447	0.0394	0.0225	0.1455	0.0260	0.0218
	$m = 0.1T$	21	0.1459	0.0517	0.0240	0.1392	0.0420	0.0211	0.1426	0.0259	0.0210
	$m = 0.1T$	51	0.1380	0.0807	0.0255	0.1304	0.0439	0.0189	0.1350	0.0260	0.0189
	$m = 0.4T$	11	0.2875	0.0206	0.0831	0.2874	0.0186	0.0829	0.2862	0.0130	0.0821
	$m = 0.4T$	21	0.2873	0.0284	0.0821	0.2859	0.0196	0.0821	0.2854	0.0129	0.0816
	$m = 0.4T$	51	0.2852	0.0441	0.0833	0.2837	0.0199	0.0809	0.2835	0.0129	0.0905

Table C.34: **Finite sample properties of XPE VI.** XPE for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	0.1115	0.1333	0.0302	0.0820	0.1470	0.0283	0.0419	0.1402	0.0214
	$m = 0.1T$	21	0.0950	0.1715	0.0384	0.0340	0.1917	0.0379	-0.0199	0.1815	0.0333
	$m = 0.1T$	51	-0.0232	0.2048	0.0425	-0.1020	0.2124	0.0555	-0.1329	0.1894	0.0535
	$m = 0.4T$	11	0.0926	0.0640	0.0127	0.0598	0.0659	0.0079	0.0180	0.0667	0.0048
	$m = 0.4T$	21	0.0948	0.0885	0.0168	0.0340	0.0885	0.0090	-0.0068	0.0833	0.0070
	$m = 0.4T$	51	0.0903	0.1420	0.0283	-0.0229	0.1329	0.0182	-0.0442	0.1186	0.0160
$T = 1000$	$m = 0.1T$	11	0.1293	0.0909	0.0250	0.0986	0.0967	0.0191	0.0583	0.1002	0.0134
	$m = 0.1T$	21	0.1268	0.1260	0.0320	0.0724	0.1307	0.0223	0.0155	0.1321	0.0177
	$m = 0.1T$	51	0.0978	0.1849	0.0437	-0.0109	0.1829	0.0336	-0.0669	0.1817	0.0375
	$m = 0.4T$	11	0.0960	0.0438	0.0111	0.0660	0.0466	0.0065	0.0229	0.0462	0.0027
	$m = 0.4T$	21	0.0983	0.0595	0.0132	0.0457	0.0635	0.0061	-0.0003	0.0579	0.0034
	$m = 0.4T$	51	0.1020	0.0944	0.0193	0.0074	0.0942	0.0089	-0.0264	0.0806	0.0072
$T = 5000$	$m = 0.1T$	11	0.1389	0.0384	0.0208	0.1139	0.0399	0.0146	0.0799	0.0421	0.0082
	$m = 0.1T$	21	0.1414	0.0524	0.0227	0.0983	0.0559	0.0127	0.0516	0.0556	0.0058
	$m = 0.1T$	51	0.1488	0.0836	0.0291	0.0639	0.0850	0.0113	0.0163	0.0806	0.0068
	$m = 0.4T$	11	0.0999	0.0181	0.0103	0.0694	0.0200	0.0052	0.0300	0.0214	0.0014
	$m = 0.4T$	21	0.1002	0.0248	0.0106	0.0493	0.0273	0.0032	0.0103	0.0261	0.0008
	$m = 0.4T$	51	0.1017	0.0391	0.0119	0.0201	0.0396	0.0020	-0.0044	0.0346	0.0012

Table C.35: **Finite sample properties of XPE VII.** XPE of H_ρ based on $\widehat{H_{xy}} - \frac{\widehat{H_x}-\widehat{H_y}}{2}$ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	0.2066	0.0673	0.0472	0.1946	0.0681	0.0425	0.1675	0.0662	0.0324
	$m = 0.1T$	21	0.2042	0.0862	0.0491	0.1800	0.0890	0.0403	0.1432	0.0850	0.0277
	$m = 0.1T$	51	0.1937	0.1052	0.0486	0.1562	0.0987	0.0342	0.1327	0.0943	0.0265
	$m = 0.4T$	11	0.2033	0.0313	0.0423	0.1868	0.0316	0.0359	0.1657	0.0290	0.0283
	$m = 0.4T$	21	0.2042	0.0419	0.0434	0.1751	0.0433	0.0325	0.1515	0.0380	0.0244
	$m = 0.4T$	51	0.2046	0.0667	0.0463	0.1518	0.0664	0.0274	0.1364	0.0560	0.0217
$T = 1000$	$m = 0.1T$	11	0.2057	0.0451	0.0443	0.1901	0.0453	0.0382	0.1684	0.0433	0.0302
	$m = 0.1T$	21	0.2042	0.0620	0.0455	0.1801	0.0615	0.0362	0.1479	0.0576	0.0252
	$m = 0.1T$	51	0.2048	0.0899	0.0500	0.1559	0.0886	0.0322	0.1193	0.0847	0.0214
	$m = 0.4T$	11	0.2022	0.0214	0.0413	0.1860	0.0213	0.0351	0.1645	0.0206	0.0275
	$m = 0.4T$	21	0.2030	0.0298	0.0421	0.1749	0.0288	0.0314	0.1526	0.0262	0.0240
	$m = 0.4T$	51	0.2045	0.0484	0.0441	0.1553	0.0435	0.0260	0.1378	0.0397	0.0206
$T = 5000$	$m = 0.1T$	11	0.2010	0.0198	0.0408	0.1901	0.0196	0.0365	0.1713	0.0189	0.0297
	$m = 0.1T$	21	0.2027	0.0259	0.0417	0.1814	0.0270	0.0336	0.1565	0.0249	0.0251
	$m = 0.1T$	51	0.2053	0.0410	0.0438	0.1639	0.0418	0.0286	0.1369	0.0371	0.0201
	$m = 0.4T$	11	0.1997	0.0094	0.0400	0.1839	0.0094	0.0339	0.1649	0.0091	0.0273
	$m = 0.4T$	21	0.1994	0.0126	0.0399	0.1733	0.0128	0.0302	0.1542	0.0114	0.0239
	$m = 0.4T$	51	0.1986	0.0200	0.0399	0.1579	0.0188	0.0253	0.1454	0.0155	0.0214

Table C.36: **Finite sample properties of XPE VIII.** XPE of H_ρ based on $\widehat{H_\rho}$ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1T$	11	0.2126	0.1181	0.0591	0.1893	0.1213	0.0505	0.1361	0.1136	0.0314
	$m = 0.1T$	21	0.2082	0.1607	0.0692	0.1608	0.1666	0.0536	0.0871	0.1578	0.0325
	$m = 0.1T$	51	0.1873	0.2055	0.0773	0.1137	0.1922	0.0499	0.0660	0.1818	0.0374
	$m = 0.4T$	11	0.2063	0.0543	0.0455	0.1715	0.0567	0.0326	0.1318	0.0508	0.0200
	$m = 0.4T$	21	0.2080	0.0783	0.0494	0.1480	0.0813	0.0285	0.1035	0.0702	0.0156
	$m = 0.4T$	51	0.2088	0.1294	0.0603	0.1017	0.1296	0.0271	0.0732	0.1072	0.0168
$T = 1000$	$m = 0.1T$	11	0.2120	0.0787	0.0511	0.1807	0.0794	0.0390	0.1411	0.0762	0.0257
	$m = 0.1T$	21	0.2089	0.1143	0.0567	0.1607	0.1147	0.0390	0.0999	0.1072	0.0215
	$m = 0.1T$	51	0.2101	0.1744	0.0745	0.1123	0.1714	0.0420	0.0423	0.1656	0.0292
	$m = 0.4T$	11	0.2045	0.0379	0.0433	0.1721	0.0385	0.0311	0.1295	0.0366	0.0181
	$m = 0.4T$	21	0.2061	0.0554	0.0455	0.1498	0.0552	0.0255	0.1059	0.0485	0.0136
	$m = 0.4T$	51	0.2090	0.0933	0.0524	0.1106	0.0856	0.0195	0.0764	0.0764	0.0117
$T = 5000$	$m = 0.1T$	11	0.2028	0.0342	0.0423	0.1811	0.0349	0.0340	0.1416	0.0334	0.0212
	$m = 0.1T$	21	0.2060	0.0479	0.0447	0.1636	0.0508	0.0294	0.1121	0.0466	0.0147
	$m = 0.1T$	51	0.2112	0.0795	0.0509	0.1290	0.0815	0.0233	0.0728	0.0715	0.0104
	$m = 0.4T$	11	0.1992	0.0166	0.0399	0.1680	0.0173	0.0285	0.1297	0.0161	0.0171
	$m = 0.4T$	21	0.1985	0.0235	0.0400	0.1468	0.0245	0.0221	0.1084	0.0211	0.0122
	$m = 0.4T$	51	0.1970	0.0391	0.0403	0.1159	0.0369	0.0148	0.0907	0.0298	0.0091

Table C.37: **Finite sample properties of LXW I.** LXW estimator for correlated ARFIMA processes with $d_1 = d_2 = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	-0.0210	0.1248	0.0160	-0.033	0.1240	0.0165	-0.0305	0.0914	0.0093
	$m = 0.1$T	21	-0.0338	0.1579	0.0261	-0.0420	0.1276	0.0180	-0.0325	0.0817	0.0077
	$m = 0.1$T	51	-0.0810	0.1695	0.0353	-0.0593	0.0771	0.0095	-0.0530	0.0477	0.0051
	$m = 0.4$T	11	-0.0127	0.0595	0.0037	-0.0198	0.0587	0.0038	-0.0150	0.0415	0.0019
	$m = 0.4$T	21	-0.0169	0.0792	0.0066	-0.0225	0.0649	0.0047	-0.0156	0.0412	0.0019
	$m = 0.4$T	51	-0.0287	0.1217	0.0156	-0.0258	0.0636	0.0047	-0.0198	0.0392	0.0019
$T = 1000$	$m = 0.1$T	11	-0.0172	0.0899	0.0084	-0.0166	0.0808	0.0068	-0.0151	0.0605	0.0039
	$m = 0.1$T	21	-0.0242	0.1166	0.0142	-0.0208	0.0869	0.0080	-0.0164	0.0583	0.0037
	$m = 0.1$T	51	-0.0482	0.1620	0.0286	-0.0285	0.0769	0.0067	-0.0261	0.0498	0.0032
	$m = 0.4$T	11	-0.0096	0.0413	0.0018	-0.0126	0.0400	0.0018	-0.0124	0.0286	0.0010
	$m = 0.4$T	21	-0.0112	0.0559	0.0033	-0.0140	0.0433	0.0021	-0.0126	0.0286	0.0010
	$m = 0.4$T	51	-0.0181	0.0883	0.0081	-0.0157	0.0446	0..0022	-0.0142	0.0282	0.0010
$T = 5000$	$m = 0.1$T	11	-0.0060	0.0352	0.0013	-0.0048	0.0352	0.0013	-0.0041	0.0259	0.0007
	$m = 0.1$T	21	-0.0095	0.0485	0.0024	-0.0052	0.0379	0.0015	-0.0042	0.0259	0.0007
	$m = 0.1$T	51	-0.0157	0.0775	0.0063	-0.0061	0.0395	0.0016	-0.0052	0.0258	0.0007
	$m = 0.4$T	11	-0.0068	0.0179	0.0004	-0.0078	0.0170	0.0003	-0.0077	0.0125	0.0002
	$m = 0.4$T	21	-0.0067	0.0237	0.0006	-0.0080	0.0180	0.0004	-0.0077	0.0126	0.0002
	$m = 0.4$T	51	-0.0079	0.0363	0.0014	-0.0081	0.0186	0.0004	-0.0078	0.0127	0.0002

Table C.38: **Finite sample properties of LXW II.** LXW estimator for correlated ARFIMA processes with $d_1 = d_2 = 0.4$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1\text{T}$	11	-0.0108	0.1390	0.0194	-0.0405	0.1396	0.0211	-0.0287	0.1025	0.0113
	$m = 0.1\text{T}$	21	-0.0158	0.1936	0.0377	-0.0574	0.1730	0.0332	-0.0332	0.1089	0.0130
	$m = 0.1\text{T}$	51	-0.1633	0.2237	0.0767	-0.1636	0.1254	0.0425	-0.1455	0.0760	0.0269
	$m = 0.4\text{T}$	11	-0.0271	0.0630	0.0047	-0.0325	0.0627	0.0050	-0.0324	0.0439	0.0030
	$m = 0.4\text{T}$	21	-0.0156	0.0863	0.0077	-0.0281	0.0733	0.0062	-0.0257	0.0476	0.0029
	$m = 0.4\text{T}$	51	0.0022	0.1444	0.0209	-0.0191	0.0921	0.0088	-0.0145	0.0560	0.0033
$T = 1000$	$m = 0.1\text{T}$	11	-0.0085	0.0943	0.0090	-0.0180	0.0867	0.0078	-0.0135	0.0686	0.0049
	$m = 0.1\text{T}$	21	-0.0017	0.1327	0.0176	-0.0193	0.1081	0.0121	-0.0068	0.0749	0.0057
	$m = 0.1\text{T}$	51	-0.0170	0.2214	0.0493	-0.0359	0.1323	0.0188	-0.0173	0.0840	0.0074
	$m = 0.4\text{T}$	11	-0.0243	0.0434	0.0025	-0.0280	0.0410	0.0025	-0.0268	0.0308	0.0017
	$m = 0.4\text{T}$	21	-0.0151	0.0590	0.0037	-0.0238	0.0474	0.0028	-0.0211	0.0328	0.0015
	$m = 0.4\text{T}$	51	0.0084	0.0956	0.0092	-0.0117	0.0596	0.0037	-0.0067	0.0382	0.0015
$T = 5000$	$m = 0.1\text{T}$	11	-0.0010	0.0373	0.0014	-0.0031	0.0352	0.0013	-0.0030	0.0260	0.0007
	$m = 0.1\text{T}$	21	0.0068	0.0518	0.0027	0.0011	0.0399	0.0016	0.0023	0.0277	0.0008
	$m = 0.1\text{T}$	51	0.0257	0.0857	0.0080	0.0139	0.0495	0.0026	0.0167	0.0324	0.0013
	$m = 0.4\text{T}$	11	-0.0244	0.0175	0.0009	-0.0249	0.0174	0.0009	-0.0251	0.0128	0.0008
	$m = 0.4\text{T}$	21	-0.0207	0.0245	0.0010	-0.0229	0.0192	0.0009	-0.0228	0.0133	0.0007
	$m = 0.4\text{T}$	51	-0.0098	0.0399	0.0017	-0.0163	0.0224	0.0008	-0.0158	0.0149	0.0005

Table C.39: **Finite sample properties of LXW III.** LXW estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.1$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	-0.0404	0.1293	0.0183	-0.0481	0.1297	0.0191	-0.0499	0.0906	0.0107
	$m = 0.1$T	21	-0.0621	0.1671	0.0318	-0.0792	0.1391	0.0256	-0.0722	0.0848	0.0124
	$m = 0.1$T	51	-0.1403	0.1851	0.0540	-0.1408	0.0953	0.0289	-0.1333	0.0510	0.0204
	$m = 0.4$T	11	0.0069	0.0612	0.0038	0.0012	0.0599	0.0036	-0.0009	0.0418	0.0017
	$m = 0.4$T	21	0.0055	0.0835	0.0070	-0.0066	0.0677	0.0046	-0.0060	0.0423	0.0018
	$m = 0.4$T	51	-0.0086	0.1307	0.0171	-0.0242	0.0750	0.0062	-0.0213	0.0424	0.0023
$T = 1000$	$m = 0.1$T	11	-0.0180	0.0885	0.0081	-0.0258	0.0874	0.0083	-0.0245	0.0663	0.0050
	$m = 0.1$T	21	-0.0250	0.1199	0.0150	-0.0425	0.0984	0.0115	-0.0358	0.0664	0.0057
	$m = 0.1$T	51	-0.0662	0.1751	0.0350	-0.0812	0.0999	0.0166	-0.0707	0.0594	0.0085
	$m = 0.4$T	11	0.0099	0.0419	0.0018	0.0040	0.0394	0.0016	0.0064	0.0283	0.0008
	$m = 0.4$T	21	0.0094	0.0568	0.0033	0.0001	0.0436	0.0019	0.0039	0.0288	0.0008
	$m = 0.4$T	51	0.0032	0.0864	0.0075	-0.0082	0.0479	0.0024	-0.0032	0.0296	0.0009
$T = 5000$	$m = 0.1$T	11	-0.0031	0.0370	0.0014	-0.0037	0.0347	0.0012	-0.0044	0.0255	0.0007
	$m = 0.1$T	21	-0.0042	0.0506	0.0026	-0.0069	0.0382	0.0015	-0.0065	0.0260	0.0007
	$m = 0.1$T	51	-0.0097	0.0807	0.0066	-0.0143	0.0422	0.0020	-0.0126	0.0269	0.0009
	$m = 0.4$T	11	0.0126	0.0181	0.0005	0.0120	0.0172	0.0004	0.0117	0.0130	0.0003
	$m = 0.4$T	21	0.0132	0.0243	0.0008	0.0114	0.0184	0.0005	0.0112	0.0131	0.0003
	$m = 0.4$T	51	0.0121	0.0383	0.0016	0.0101	0.0195	0.0005	0.0101	0.0132	0.0003

Table C.40: **Finite sample properties of LXW IV.** LXW estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.5$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.0168	0.1371	0.0191	-0.0111	0.1363	0.0187	-0.0057	0.0900	0.0081
	$m = 0.1$T	21	-0.0026	0.1796	0.0323	-0.0405	0.1496	0.0240	-0.0261	0.0852	0.0079
	$m = 0.1$T	51	-0.1030	0.1932	0.0480	-0.0970	0.0962	0.0187	-0.0870	0.0503	0.0101
	$m = 0.4$T	11	0.1542	0.0611	0.0275	0.1486	0.0612	0.0258	0.1506	0.0453	0.0247
	$m = 0.4$T	21	0.1544	0.0844	0.0309	0.1444	0.0672	0.0254	0.1479	0.0460	0.0240
	$m = 0.4$T	51	0.1450	0.1364	0.0396	0.1327	0.0769	0.0235	0.1384	0.0474	0.0214
$T = 1000$	$m = 0.1$T	11	0.0235	0.0881	0.0083	0.0167	0.0912	0.0086	0.0146	0.0656	0.0045
	$m = 0.1$T	21	0.0156	0.1212	0.0149	0.0028	0.1024	0.0105	0.0044	0.0658	0.0043
	$m = 0.1$T	51	-0.0225	0.1796	0.0327	-0.0315	0.1039	0.0118	-0.0269	0.0598	0.0043
	$m = 0.4$T	11	0.1538	0.0435	0.0256	0.1523	0.0409	0.0249	0.1527	0.0304	0.0242
	$m = 0.4$T	21	0.1532	0.0590	0.0270	0.1506	0.0447	0.0247	0.1517	0.0308	0.0240
	$m = 0.4$T	51	0.1531	0.0938	0.0322	0.1463	0.0495	0.0239	0.1491	0.0321	0.0232
$T = 5000$	$m = 0.1$T	11	0.0337	0.0380	0.0026	0.0277	0.0350	0.0020	0.0306	0.0254	0.0016
	$m = 0.1$T	21	0.0330	0.0517	0.0038	0.0251	0.0385	0.0021	0.0289	0.0257	0.0015
	$m = 0.1$T	51	0.0300	0.0793	0.0072	0.0187	0.0430	0.0022	0.0239	0.0266	0.0013
	$m = 0.4$T	11	0.1524	0.0191	0.0236	0.1524	0.0179	0.0235	0.1524	0.0132	0.0234
	$m = 0.4$T	21	0.1526	0.0258	0.0240	0.1523	0.0191	0.0236	0.1524	0.0132	0.0234
	$m = 0.4$T	51	0.1549	0.0401	0.0256	0.1522	0.0199	0.0236	0.1525	0.0134	0.0234

Table C.41: **Finite sample properties of LXW V.** LXW estimator for correlated ARFIMA and AR(1) processes with $d = 0.4$ and $\theta = 0.8$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.1573	0.1368	0.0435	0.1364	0.1401	0.0382	0.1422	0.0998	0.0302
	$m = 0.1$T	21	0.1349	0.1800	0.0506	0.1083	0.1556	0.0359	0.1225	0.0973	0.0245
	$m = 0.1$T	51	-0.0206	0.2073	0.0434	0.0136	0.1001	0.0102	0.0244	0.0594	0.0041
	$m = 0.4$T	11	0.3173	0.0636	0.1047	0.3134	0.0634	0.1022	0.3170	0.0447	0.1025
	$m = 0.4$T	21	0.3225	0.0857	0.1114	0.3132	0.0708	0.1031	0.3184	0.0460	0.1035
	$m = 0.4$T	51	0.3374	0.1406	0.1336	0.3168	0.0867	0.1079	0.3256	0.0516	0.1087
$T = 1000$	$m = 0.1$T	11	0.1619	0.0921	0.0347	0.1518	0.0879	0.0308	0.1551	0.0663	0.0285
	$m = 0.1$T	21	0.1554	0.1271	0.0403	0.1390	0.1010	0.0295	0.1476	0.0678	0.0264
	$m = 0.1$T	51	0.1096	0.1965	0.0506	0.1040	0.1128	0.0235	0.1170	0.0653	0.0180
	$m = 0.4$T	11	0.3125	0.0451	0.0997	0.3111	0.0423	0.0986	0.3115	0.0309	0.0980
	$m = 0.4$T	21	0.3142	0.0604	0.1024	0.3114	0.0455	0.0991	0.3123	0.0310	0.0985
	$m = 0.4$T	51	0.3242	0.0941	0.1140	0.3153	0.0512	0.1021	0.3173	0.0324	0.1017
$T = 5000$	$m = 0.1$T	11	0.1600	0.0394	0.0272	0.1602	0.0364	0.0270	0.1591	0.0269	0.0260
	$m = 0.1$T	21	0.1603	0.0529	0.0285	0.1594	0.0391	0.0269	0.1585	0.0272	0.0259
	$m = 0.1$T	51	0.1607	0.0804	0.0323	0.1569	0.0424	0.0264	0.1569	0.0281	0.0254
	$m = 0.4$T	11	0.3085	0.0195	0.0955	0.3079	0.0187	0.0952	0.3086	0.0135	0.0954
	$m = 0.4$T	21	0.3090	0.0261	0.0962	0.3081	0.0197	0.0953	0.3088	0.0135	0.0955
	$m = 0.4$T	51	0.3112	0.0404	0.0985	0.3090	0.0204	0.0959	0.3096	0.0136	0.0960

Table C.42: **Finite sample properties of LXW VI.** LXW estimator for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1\text{T}$	11	0.1145	0.1363	0.0317	0.0875	0.1349	0.0259	0.0442	0.1398	0.0215
	$m = 0.1\text{T}$	21	0.1065	0.1866	0.0461	0.0514	0.1865	0.0374	-0.0145	0.1810	0.0330
	$m = 0.1\text{T}$	51	-0.0247	0.2256	0.0515	-0.0888	0.2082	0.0512	-0.1293	0.1773	0.0481
	$m = 0.4\text{T}$	11	0.0962	0.0600	0.0128	0.0582	0.0632	0.0074	0.0230	0.0611	0.0043
	$m = 0.4\text{T}$	21	0.1024	0.0815	0.0171	0.0405	0.0835	0.0086	-0.0007	0.0777	0.0060
	$m = 0.4\text{T}$	51	0.1090	0.1359	0.0304	-0.0092	0.1300	0.0170	-0.0412	0.1145	0.0148
$T = 1000$	$m = 0.1\text{T}$	11	0.1268	0.0887	0.0239	0.1003	0.0901	0.0182	0.0680	0.0914	0.0130
	$m = 0.1\text{T}$	21	0.1325	0.1215	0.0323	0.0828	0.1275	0.0231	0.0327	0.1239	0.0164
	$m = 0.1\text{T}$	51	0.1182	0.1964	0.0525	0.0108	0.1984	0.0395	-0.0533	0.1839	0.0366
	$m = 0.4\text{T}$	11	0.0996	0.0405	0.0116	0.0669	0.0431	0.0063	0.0332	0.0422	0.0029
	$m = 0.4\text{T}$	21	0.1042	0.0564	0.0140	0.0515	0.0585	0.0061	0.0139	0.0538	0.0031
	$m = 0.4\text{T}$	51	0.1142	0.0928	0.0217	0.0200	0.0881	0.0082	-0.0138	0.0765	0.0060
$T = 5000$	$m = 0.1\text{T}$	11	0.1406	0.0365	0.0211	0.1183	0.0390	0.0155	0.0846	0.0367	0.0085
	$m = 0.1\text{T}$	21	0.1458	0.0514	0.0239	0.1087	0.0526	0.0146	0.0632	0.0490	0.0064
	$m = 0.1\text{T}$	51	0.1629	0.0838	0.0335	0.0872	0.0825	0.0144	0.0323	0.0744	0.0066
	$m = 0.4\text{T}$	11	0.1018	0.0182	0.0107	0.0735	0.0188	0.0058	0.0405	0.0187	0.0020
	$m = 0.4\text{T}$	21	0.1041	0.0253	0.0115	0.0593	0.0253	0.0042	0.0241	0.0235	0.0011
	$m = 0.4\text{T}$	51	0.1102	0.0400	0.0138	0.0385	0.0383	0.0029	0.0081	0.0318	0.0011

Table C.43: **Finite sample properties of LXW VII.** LXW estimator of H_ρ based on $\widehat{H_{xy}} - \frac{\widehat{H_x} - \widehat{H_y}}{2}$ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1$T	11	0.2120	0.1085	0.0567	0.1863	0.1083	0.0464	0.1381	0.1099	0.0311
	$m = 0.1$T	21	0.2124	0.1576	0.0700	0.1592	0.1604	0.0511	0.0895	0.1583	0.0331
	$m = 0.1$T	51	0.1854	0.2111	0.0789	0.1104	0.1944	0.0500	0.0672	0.1797	0.0368
	$m = 0.4$T	11	0.2043	0.0506	0.0443	0.1685	0.0515	0.0311	0.1331	0.0466	0.0199
	$m = 0.4$T	21	0.2041	0.0744	0.0472	0.1480	0.0734	0.0273	0.1058	0.0637	0.0152
	$m = 0.4$T	51	0.2055	0.1267	0.0583	0.0994	0.1170	0.0236	0.0613	0.1052	0.0148
$T = 1000$	$m = 0.1$T	11	0.2058	0.0752	0.0480	0.1817	0.0748	0.0386	0.1415	0.0689	0.0248
	$m = 0.1$T	21	0.2096	0.1113	0.0563	0.1631	0.1099	0.0387	0.1062	0.0963	0.0206
	$m = 0.1$T	51	0.2186	0.1968	0.0865	0.1098	0.1849	0.0463	0.0411	0.1656	0.0291
	$m = 0.4$T	11	0.1991	0.0347	0.0409	0.1725	0.0352	0.0310	0.1369	0.0314	0.0197
	$m = 0.4$T	21	0.1986	0.0512	0.0421	0.1544	0.0507	0.0264	0.1137	0.0433	0.0148
	$m = 0.4$T	51	0.2012	0.0894	0.0485	0.1170	0.0804	0.0202	0.0881	0.0683	0.0106
$T = 5000$	$m = 0.1$T	11	0.2004	0.0304	0.0411	0.1786	0.0326	0.0330	0.1466	0.0311	0.0225
	$m = 0.1$T	21	0.2021	0.0453	0.0429	0.1639	0.0467	0.0290	0.1219	0.0439	0.0168
	$m = 0.1$T	51	0.2118	0.0773	0.0508	0.1319	0.0766	0.0233	0.0815	0.0679	0.0112
	$m = 0.4$T	11	0.1998	0.0155	0.0402	0.1704	0.0161	0.0293	0.1400	0.0143	0.0198
	$m = 0.4$T	21	0.2000	0.0219	0.0405	0.1539	0.0231	0.0242	0.1219	0.0196	0.0152
	$m = 0.4$T	51	0.2010	0.0359	0.0417	0.1275	0.0358	0.0175	0.1007	0.0286	0.0110

Table C.44: **Finite sample properties of LXW VIII.** LXW estimator of H_ρ based on $\widehat{H_\rho}$ for Mixed-correlated ARFIMA processes with $d_1 = d_4 = 0.4$, $d_2 = d_3 = 0.2$ and varying $\rho_{\varepsilon\nu}$.

		smoothing	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.9$		
			bias	SD	MSE	bias	SD	MSE	bias	SD	MSE
$T = 500$	$m = 0.1\text{T}$	11	0.2050	0.0993	0.0519	0.1844	0.0982	0.0437	0.1424	0.0980	0.0299
	$m = 0.1\text{T}$	21	0.2015	0.1419	0.0607	0.1567	0.1431	0.0450	0.0989	0.1411	0.0297
	$m = 0.1\text{T}$	51	0.1807	0.1881	0.0681	0.1179	0.1719	0.0435	0.0842	0.1556	0.0313
	$m = 0.4\text{T}$	11	0.2026	0.0457	0.0431	0.1731	0.0455	0.0320	0.1428	0.0416	0.0221
	$m = 0.4\text{T}$	21	0.2015	0.0691	0.0454	0.1527	0.0670	0.0278	0.1160	0.0580	0.0168
	$m = 0.4\text{T}$	51	0.2001	0.1143	0.0531	0.1120	0.1029	0.0231	0.0806	0.0897	0.0145
$T = 1000$	$m = 0.1\text{T}$	11	0.2022	0.0674	0.0454	0.1821	0.0668	0.0376	0.1492	0.0618	0.0261
	$m = 0.1\text{T}$	21	0.2029	0.1016	0.0515	0.1635	0.0989	0.0365	0.1157	0.0868	0.0209
	$m = 0.1\text{T}$	51	0.2061	0.1707	0.0716	0.1196	0.1561	0.0387	0.0687	0.1370	0.0235
	$m = 0.4\text{T}$	11	0.1991	0.0312	0.0406	0.1768	0.0313	0.0322	0.1460	0.0279	0.0221
	$m = 0.4\text{T}$	21	0.1979	0.0475	0.0414	0.1601	0.0460	0.0277	0.1234	0.0393	0.0168
	$m = 0.4\text{T}$	51	0.1988	0.0811	0.0461	0.1282	0.0719	0.0216	0.0922	0.0586	0.0119
$T = 5000$	$m = 0.1\text{T}$	11	0.2002	0.0270	0.0408	0.1818	0.0286	0.0339	0.1556	0.0272	0.0250
	$m = 0.1\text{T}$	21	0.2012	0.0419	0.0423	0.1679	0.0428	0.0300	0.1315	0.0401	0.0189
	$m = 0.1\text{T}$	51	0.2084	0.0711	0.0485	0.1395	0.0695	0.0243	0.0960	0.0606	0.0129
	$m = 0.4\text{T}$	11	0.1996	0.0136	0.0400	0.1747	0.0140	0.0307	0.1488	0.0123	0.0223
	$m = 0.4\text{T}$	21	0.1995	0.0199	0.0402	0.1593	0.0204	0.0258	0.1308	0.0170	0.0174
	$m = 0.4\text{T}$	51	0.1996	0.0328	0.0409	0.1356	0.0319	0.0194	0.1106	0.0243	0.0128

Appendix D

Tables and figures for Chapter 6

	ADF_1	ADF_2	PP	$KPSS_1$	$KPSS_2$
AEX	-16.6222***	-16.6454***	-3519.977***	0.1589	0.1016
BVSPA	-27.1395***	-27.1360***	-2868.923***	0.1279	0.1270*
CAC	-20.2046***	-20.2173***	-3273.002***	0.1643	0.1298*
DAX	-16.9515***	-17.0795***	-3192.172***	0.5218**	0.1338*
DJIA	-23.6831***	-23.6987***	-3279.574***	0.0962	0.0378
EuroSTOXX	-20.4521***	-20.4599***	-3181.928***	0.1314	0.0910
FTSE	-19.4146***	-19.4579***	-3231.826***	0.1872	0.0569
HSI	-18.0427***	-18.0412***	-2838.066***	0.1870	0.1888
IBEX	-18.7168***	-18.7140***	-3263.749***	0.1524	0.1560**
KOSPI	-23.7026***	-23.6997***	-3090.337***	0.0454	0.0427
NASDAQ	-17.3697***	-17.5840***	-3218.774***	0.7251**	0.1226*
NIKKEI	-20.3690***	-20.3939***	-3025.452***	0.1137	0.049
SPX	-23.6049***	-23.6445***	-3253.274***	0.2090	0.0585
SSMI	-27.1479***	-27.1559***	-3125.052***	0.1022	0.0808

Table D.1: **Unit-root and stationarity tests – returns.** Augmented Dickey-Fuller tests (ADF) are run with (ADF_1) and without (ADF_2) a constant term with lags based on Akaike information criterion (AIC) with a maximum lag of 10. KPSS test are used with a constant ($KPSS_1$) and with a constant and a time trend ($KPSS_2$) with the same lag selection as for ADF. ADF and PP tests have a null hypothesis of a unit root process, KPSS test has a null of a stationary process.

	ADF_1	ADF_2	PP	$KPSS_1$	$KPSS_2$
AEX	-4.8205***	-4.8519***	-404.5102***	1.7032***	1.5622***
BVSPA	-7.6172***	-7.6172***	-1267.851***	0.6063**	0.6139***
CAC	-5.0766***	-5.0783***	-421.4923***	2.3125***	2.3338***
DAX	-4.7500***	-4.9121***	-441.1234***	3.2686***	1.8857***
DJIA	-5.3328***	-5.4726***	-898.6795***	1.5492***	1.4261***
EuroSTOXX	-5.1974***	-5.1966***	-612.0555***	1.7441***	1.7475***
FTSE	-4.5427***	-4.5776***	-425.7603***	2.1643***	2.0962***
HSI	-4.7808***	-5.1138***	-1035.481***	3.9635***	1.8088***
IBEX	-4.5560***	-4.6397***	-358.1039***	4.6870***	3.6969***
KOSPI	-5.3960***	-6.5683***	-783.4318***	8.9232***	0.9079***
NASDAQ	-4.6472***	-5.1769***	-503.5288***	6.8655***	2.4132***
NIKKEI	-5.8642***	-6.6042***	-843.0256***	4.4692***	1.0513***
SPX	-5.3373***	-5.2345***	-743.2767***	1.676***	1.6318***
SSMI	-4.5859***	-4.5931***	-353.8441***	1.0995***	1.1143***

Table D.2: **Unit-root and stationarity tests – logarithmic realized volatility.** Augmented Dickey-Fuller tests (ADF) are run with (ADF_1) and without (ADF_2) a constant term with lags based on Akaike information criterion (AIC) with a maximum lag of 10. KPSS test are used with a constant ($KPSS_1$) and with a constant and a time trend ($KPSS_2$) with the same lag selection as for ADF. ADF and PP tests have a null hypothesis of a unit root process, KPSS test has a null of a stationary process.

	ADF_1	ADF_2	PP	$KPSS_1$	$KPSS_2$
AEX	-17.1824***	-17.2573***	-3334.682***	0.3200	0.1280*
BVSPA	-26.6256***	-26.6235***	-2966.64***	0.3638*	0.3603***
CAC	-21.314***	-21.3505***	-3226.182***	0.3073	0.2177**
DAX	-17.1168***	-17.2821***	-3182.164***	0.7968***	0.3352***
DJIA	-26.3101***	-26.4039***	-3267.516***	0.5016**	0.0622
EuroSTOXX	-26.056***	-26.0655***	-3120.762***	0.2577	0.2136**
FTSE	-35.3912***	-25.3668***	-3111.298***	0.3224	0.0348
HSI	-26.4295***	-26.4707***	-2852.553***	0.5345**	0.3660***
IBEX	-58.3557***	-58.3468***	-3296.907***	0.3504*	0.3482***
KOSPI	-31.3284***	-31.326-***	-3104.152***	0.1031	0.0963
NASDAQ	-41.9536***	-42.1667***	-3155.367***	1.2649***	0.0533
NIKKEI	-20.2143***	-20.2583***	-3146.528***	0.2357	0.0991
SPX	-26.9363***	-27.1181***	-3163.449***	0.8962***	0.0957
SSMI	-57.9141***	-26.3243***	-3166.386***	0.2208	0.1730**

Table D.3: **Unit-root and stationarity tests – standardized returns.** Augmented Dickey-Fuller tests (ADF) are run with (ADF_1) and without (ADF_2) a constant term with lags based on Akaike information criterion (AIC) with a maximum lag of 10. KPSS test are used with a constant ($KPSS_1$) and with a constant and a time trend ($KPSS_2$) with the same lag selection as for ADF. ADF and PP tests have a null hypothesis of a unit root process, KPSS test has a null of a stationary process.

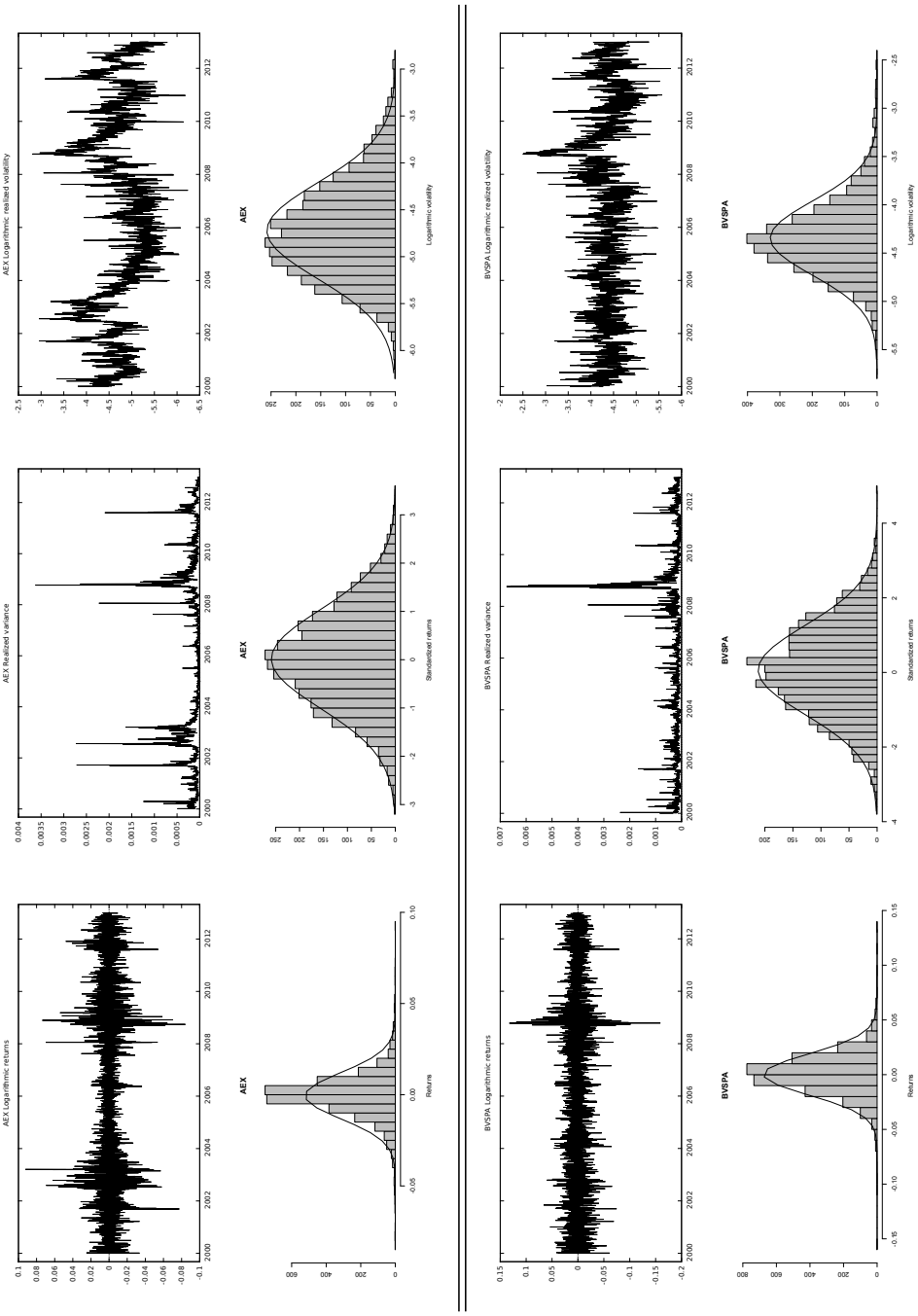


Figure D.1: AEX & BVSPA

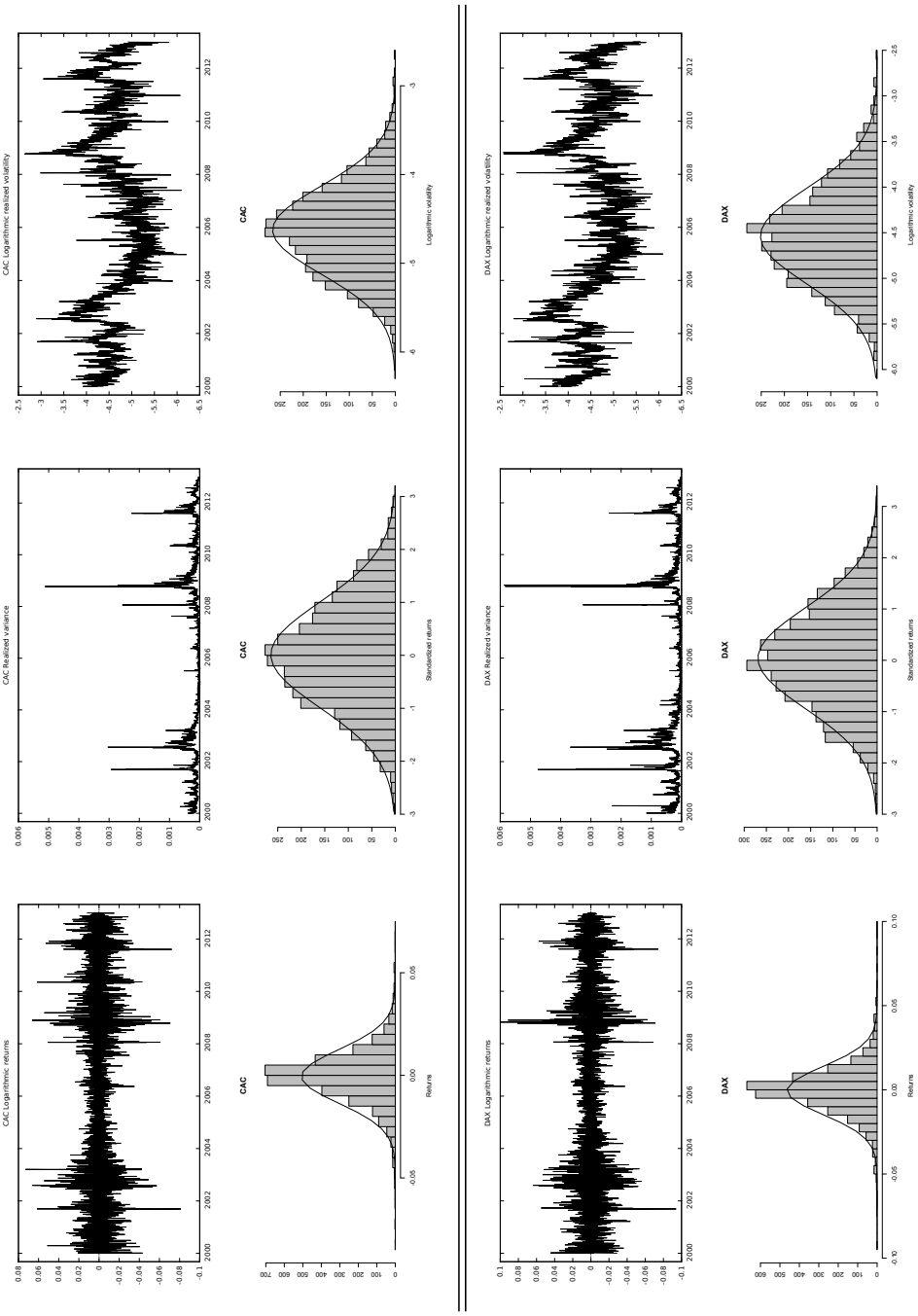


Figure D.2: CAC & DAX

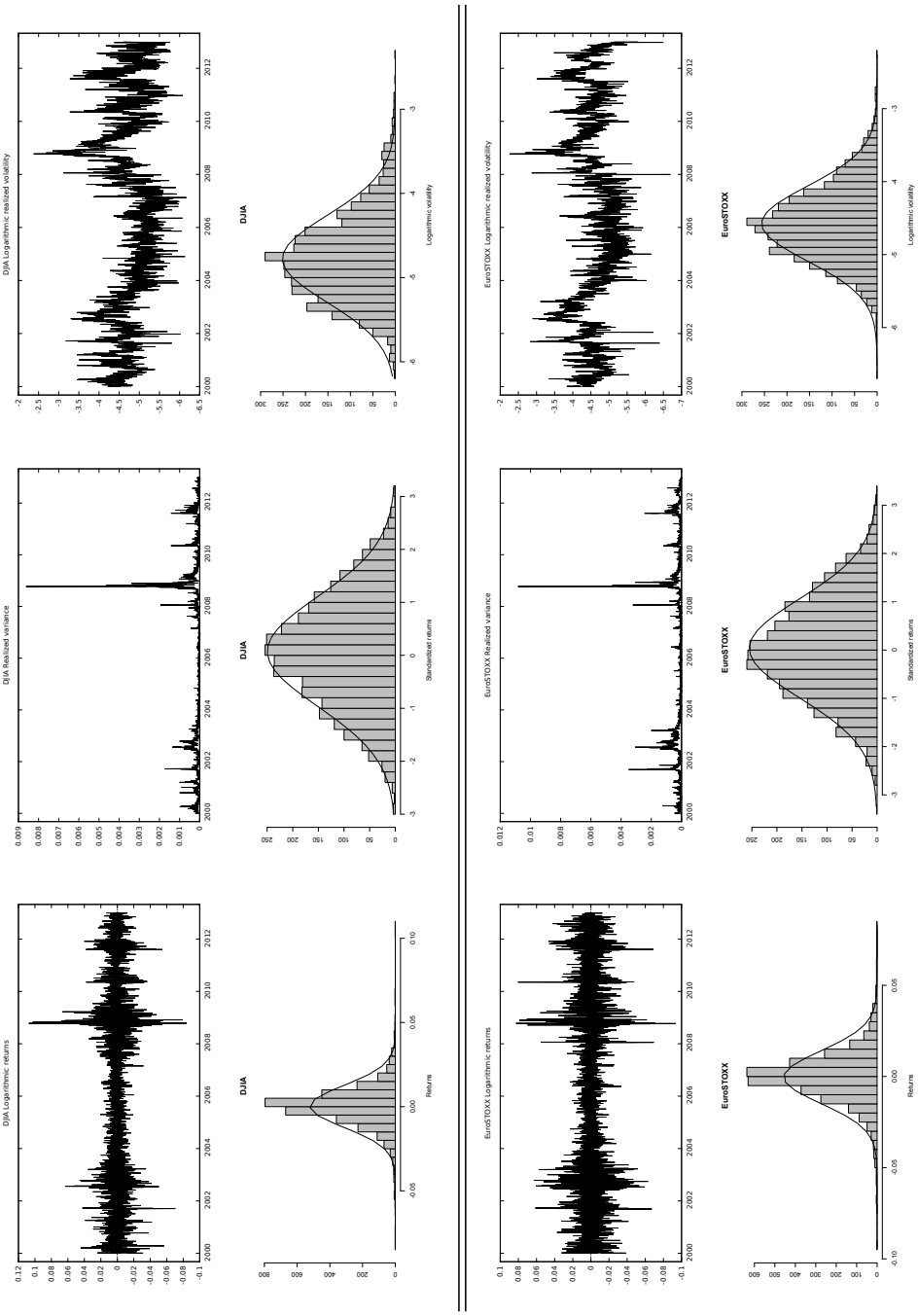


Figure D.3: DJIA & EuroSTOXX

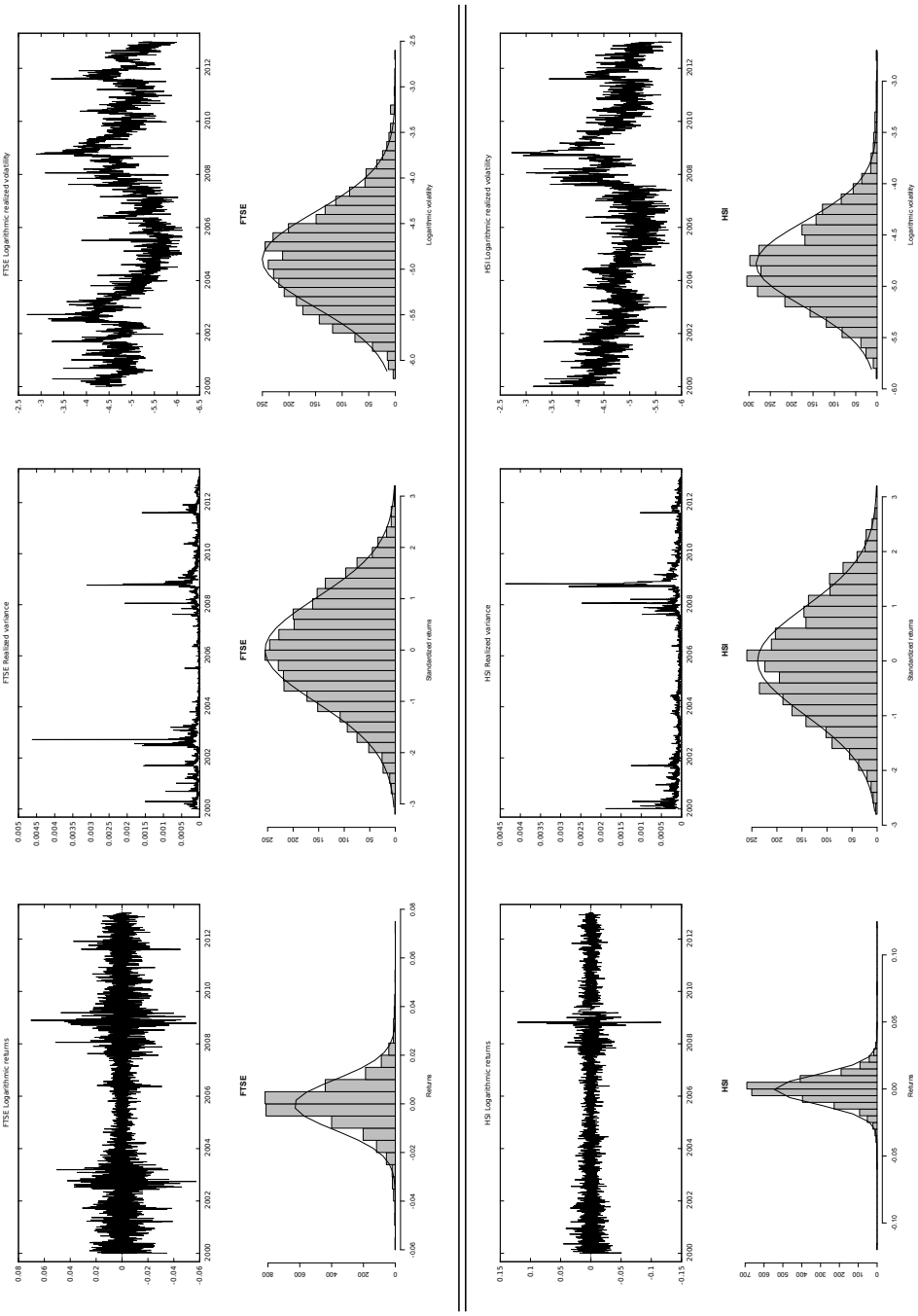


Figure D.4: FTSE & HSI

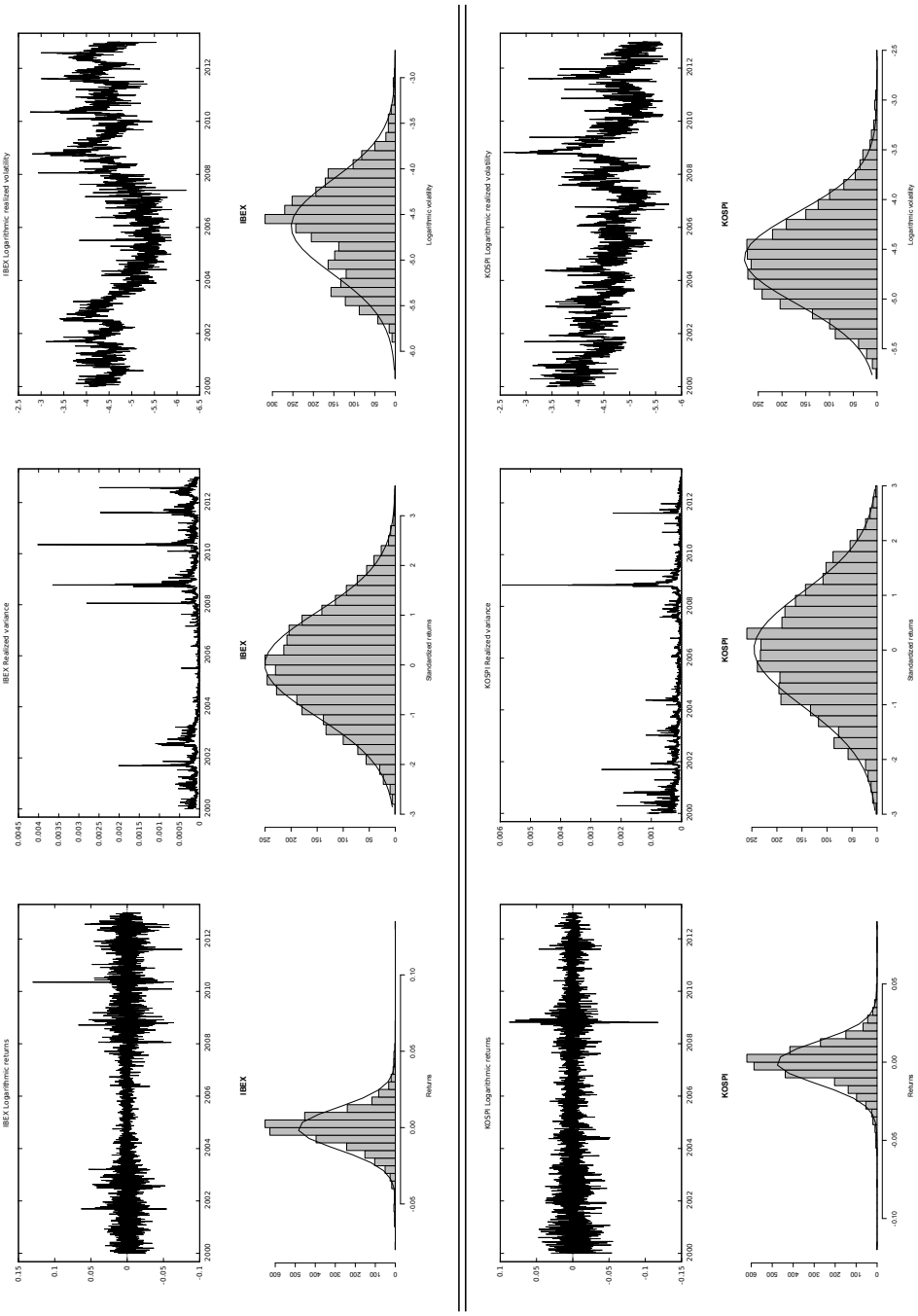
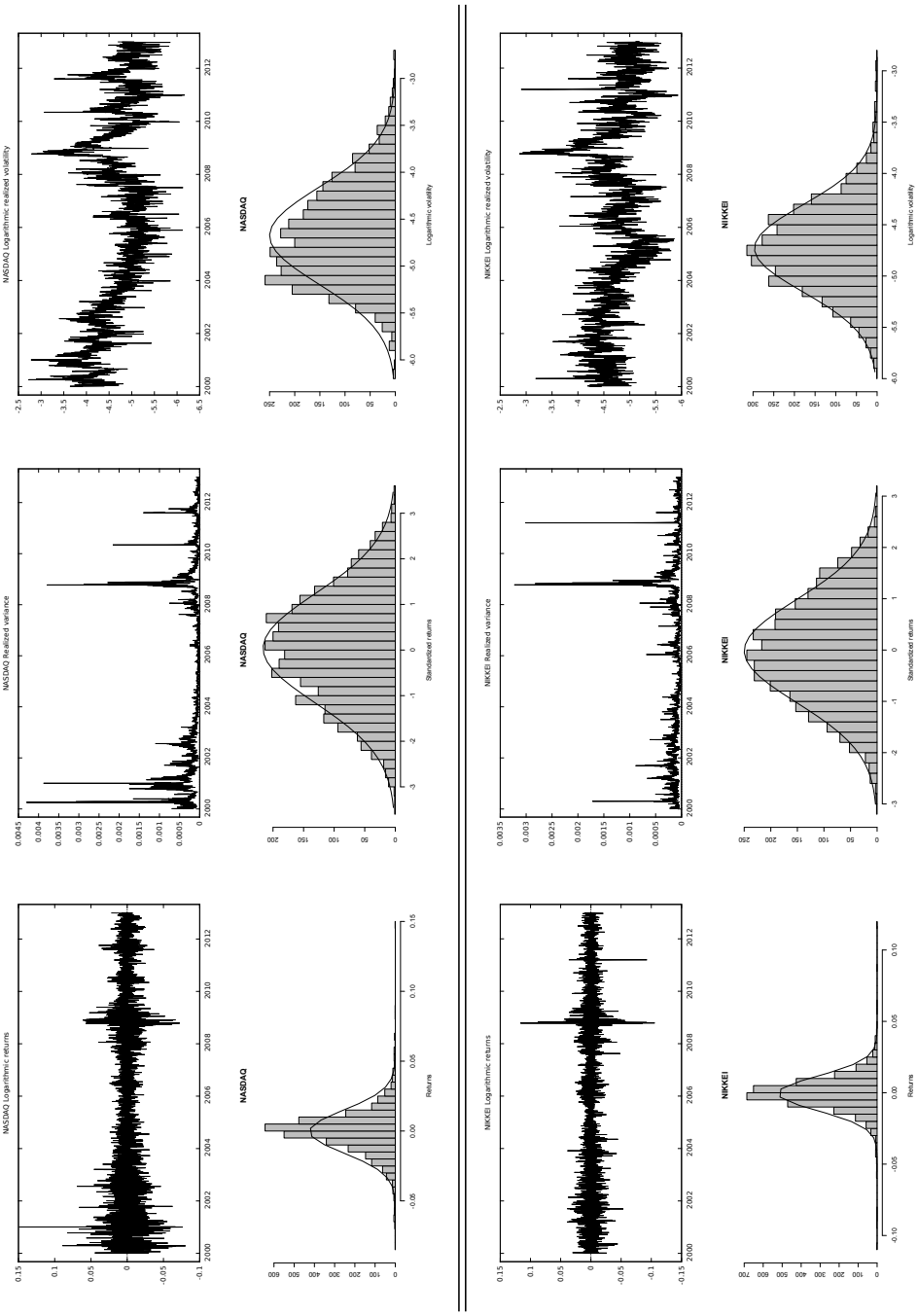


Figure D.5: IBEX & KOSPI

Figure D.6: NASD & NIKKEI



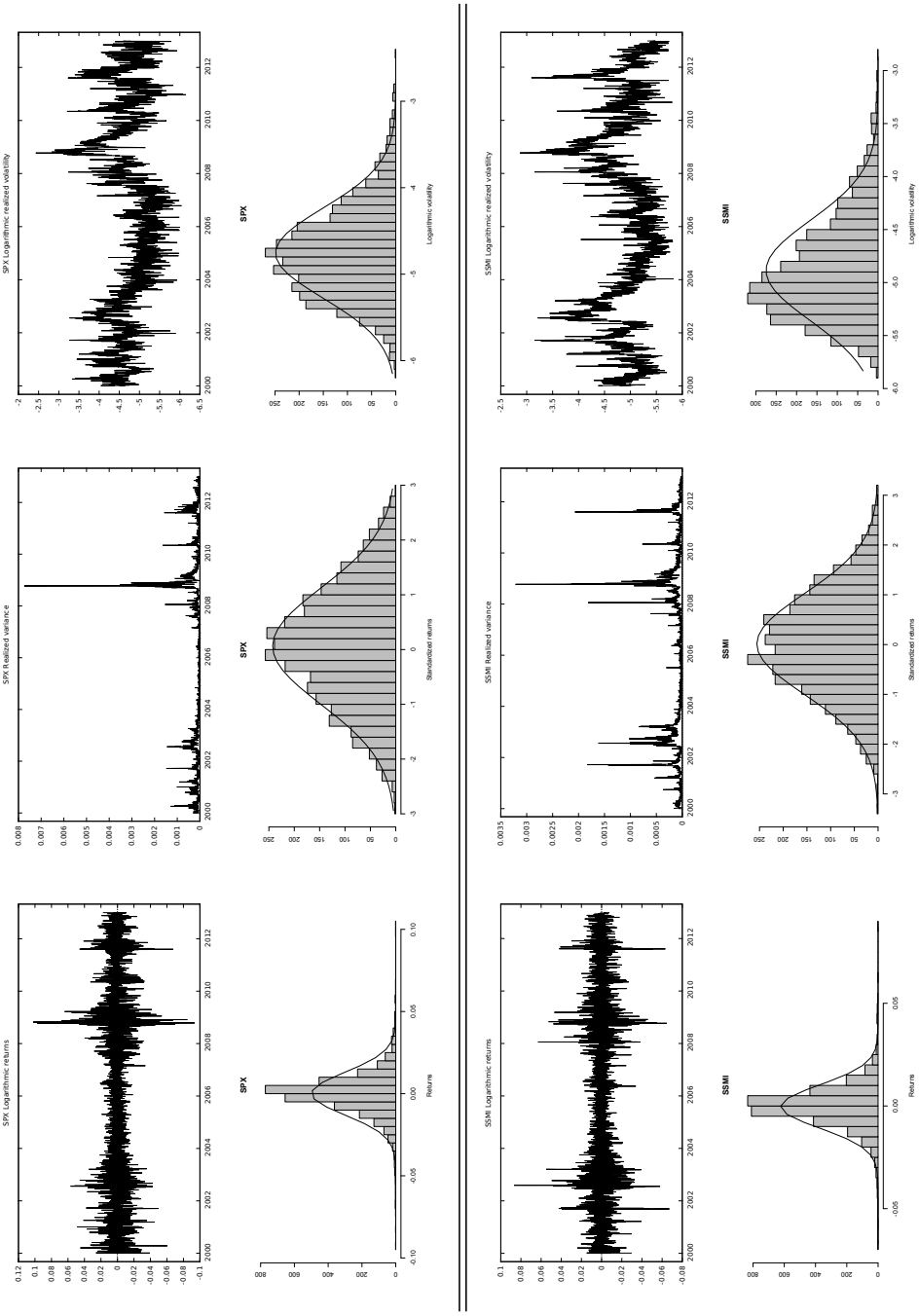


Figure D.7: SPX & SSMI